2010

Type 2 Diabetes in Children: Estimates of Epidemiology, Quality of Care, Costs and Resource Utilization in a Medicaid Population

Vennela Thumula

AN EVALUATION OF MEDICAID PROGRAMS USING MEDICATION USE-RELATED QUALITY MEASURES

A Dissertation
presented in partial fulfillment of requirements
for the degree of Doctor of Philosophy
in the Department of Pharmacy Administration
The University of Mississippi

VENNELA THUMULA

August 2012

ABSTRACT

Objective: To assess quality of care provided to Medicaid beneficiaries using medication use-related quality indicators.

Methods: This study is a retrospective analysis of 2006 & 2007 Medicaid administrative claims data from 45 states and DC. Eleven medication use-related measures (adherence/persistence and standard of care) from the AHRQ initial core set were included. A composite measure of medication use-related quality was also created. Patient case-mix adjusted measure scores were computed using hierarchical logistic regression models. States were ranked on both case-mix adjusted and unadjusted scores and categorized into top (~20%), medium (60%) and bottom (~20%) performers. Agreement in rankings and groups based on adjusted and unadjusted scores were determined using Kendall's τ_b and Cohen's κ. Cross-state variations in measure scores were described using coefficient of variation and choropleth maps. Multi-level models were used to assess the amount of variation in measures explained by the state level.

Results: National benchmarks on medication use-related measures for Medicaid for 2007 ranged from 31.5% for the AD chronic measure to 66.8% for the ICS measure. There was substantial variation in the 13 measures being studied with coefficient of variation ranging from 6.7 for the ICS measure to 20.5 for the MI1 measure. The best performing state Medicaid programs also had significant room for improvement across all measures. There was a lack of agreement in grouping based on crude and case-mix adjusted methods for majority of the

measures (κ=0.22–0.74), except for the ICS measure (κ=0.91). A very small proportion of variation in the study measures (1.5 – 5.7%) was explained by the state level random effect.

Conclusions: This study highlights the need for including medication use-related measures in the Medicaid adult quality measure set, considering the substantial variation in scores across states and the considerable room for improvement. States could create a composite measure of medication use-related quality using the approach used in this study if they find it burdensome to report on multiple measures. The study showcased the lack of agreement in crude and case-mix adjusted scores. Medicaid programs should consider the study findings before publicly reporting on crude scores.

DEDICATION

to mom

LIST OF ABBREVIATIONS AND SYMBOLS

ACEI	Angiotensin-Converting-Enzyme Inhibitor
AD	Antidepressant
AHRQ	Agency for Healthcare Research & Quality
ARB	Angiotensin Receptor Blocker
BB	Beta Blocker
BIGU	Biguanide
CAD	Coronary Artery Disease
CCB	Calcium Channel Blocker
CCI	Charlson's Comorbidity Index
CV	Coefficient of Variation
CMS	Centers for Medicare & Medicaid Services
HHS	The United States Department of Health and Human Services
ICC	Intraclass Correlation Coefficient
ICD-9-CM	International Classification of Disease, 9^{th} Revision, Clinical Modification
ICS	Inhaled Corticosteroid
LABA	Long-Acting Beta-2 Adrenergic Agonist
MAX	Medicaid Analytic Extract
MI	Myocardial Infarction

NCQA	National Committee for Quality Assurance
NDC	National Drug Code
PDC	Proportion of Days Covered
PQA	Pharmacy Quality Alliance
ResDAC	Research Data Assistance Center
SD	Standard Deviation
STAT	Statin
SU	Sulfonylurea
TZD	Thiazolidinedione
κ	Cohen's Kappa
τ	Kendall's Tau

ACKNOWLEDGMENTS

First and foremost, I owe my deepest gratitude to my advisor, Ben Banahan, for encouraging me to pursue a doctoral degree and for his phenomenal support, enthusiasm and guidance ever since. He allowed me to pursue a question that required significant time and resources, and has generously provided both. I will be forever grateful to him for his invaluable contribution to this dissertation and for his advice on several other occasions. My dissertation committee members, John Bentley, Yi Yang, Donna West Strum, and Sarahmona Przybyla have been very helpful in my timely completion of this dissertation. Their thoughtful suggestions have definitely refined my work.

I would like to thank my fellow graduate students Su Bunniran, Tasneem Lokhandwala, Clive Mendonca, and Kyle Null and my loving friends outside the department for the fun, help and encouragement through the stressful times in graduate school. I am also grateful to my friend Sai Dharmarajan for helping me with multilevel modeling.

Most importantly, I would like to thank my mom, for her unconditional love, inspiration, constant support and encouragement through everything I pursued, including graduate school. To her I dedicate this dissertation. My dad, my little sister Nishu, and my grandparents for always being there for me. I cannot finish without saying how grateful I am to my boyfriend, Saral Kal for his love, encouragement and patience through the good times and bad. I love you all with all my heart.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER I - INTRODUCTION

Quality improvement is an issue of fundamental importance in healthcare in the United States. The last two decades has seen an increase in quality measurement efforts, benchmarking and quality reporting initiatives at the provider, facility and healthcare system levels.[1-5] An important target of quality improvement efforts is the management of chronic conditions.[5-9] Numerous quality indicators used to assess provider quality are related to medication safety and effectiveness in the management of chronic conditions.[3,5,10,11] These quality indicators can be applied to state Medicaid programs to assess the quality of chronic care provided to beneficiaries. The overarching goal of this study is to assess the quality of chronic care provided by Medicaid using medication-related chronic care quality indicators proposed by the Agency for Healthcare Research and Quality (AHRQ).

On January 1, 2011 AHRQ published an initial set of quality indicators that can be used by state Medicaid programs to assess healthcare quality, as required by the Affordable Care Act.[5] This study used chronic care quality indicators related to medication use compiled by AHRQ to assess the quality of state Medicaid programs. Since, state Medicaid programs will soon adopt these measures for quality reporting, this study aims to assess the feasibility of using Medicaid data to measure quality of chronic care provided to beneficiaries. This study will also provide national estimates of chronic care quality, adjusting for patient case-mix, which will enable benchmarking performance of individual state Medicaid programs against national averages.

State Medicaid programs vary widely in terms of their eligibility requirements and program generosity.[12,13] These differences across states policies may result in variation in chronic care quality. This study will provide insights into the geographic variation in chronic care quality across the state Medicaid programs. Additionally, a substantial amount of variation in chronic care quality should be attributed to the state Medicaid level to justify quality improvement efforts at this level. Therefore, this study will assess the proportion of variability in chronic care quality explained at the patient and the state Medicaid levels using multi-level models.

The specific objectives of this study are:

1. To assess the feasibility of using state Medicaid data to provide national estimates of chronic care quality on medication use-related metrics
2. To estimate and compare crude and case-mix adjusted scores of chronic care quality on medication use-related indicators for state Medicaid programs
3. To describe the distribution of state Medicaid performance scores of chronic care quality on medication use-related measures
4. To illustrate cross-state variation in performance scores of chronic care quality on medication use-related measures
5. To assess the variation in chronic care quality that may be attributed to the patient and state Medicaid levels

There are no firm hypotheses for this study. Rather, this study should be viewed as hypotheses generating. Nevertheless, on the basis of prior research, this study anticipated that (a)

Medicaid data can be used to assess chronic care quality indicators for a majority of states, (b) case-mix adjusted scores will be different from crude estimates of chronic care quality, (c) chronic care quality scores will vary across states, and (d) patients will explain the most variance in chronic care quality scores, followed by Medicaid programs.

This study is a retrospective analysis of Medicaid administrative claims data from 45 states with at least 100,000 enrollees in 2007, and the District of Columbia. Eleven measures from the AHRQ initial core set of measures were used to assess chronic care quality. All quality scores were aggregated to the state level to provide crude estimates of chronic care quality. Additionally, a composite measure of medication use-related quality was created at the state level. Performance scores were adjusted for patient case-mix and the distribution of states on adjusted scores was reported. Cross-state variations in chronic care performance scores were described. Finally, multi-level models were used to explain which level (patient and state Medicaid) contributes the most to variation in chronic care quality.

This is the first study to apply AHRQ measures to study chronic care quality. This study proposed to compute indicators for all 45 states and provide national benchmarks for chronic care quality. The results will provide insights into the quality of chronic care provided by the 45 state Medicaid programs. This study will allow individual state Medicaid programs to compare directly with other states in terms of quality of chronic care. It will also give an assessment of the variation in chronic care quality within individual states. Additionally, this study will inform researchers and state Medicaid program officials about the extent to which they can assess chronic care quality using Medicaid datasets.

CHAPTER II – LITERATURE REVIEW

This chapter contains a brief overview of the importance of quality measurement and existing quality metrics, followed by a review of the AHRQ chronic care quality measures related to medication use and empirical evidence to judge their usability as reliable quality indicators. This section also outlines the need for case-mix adjustment and provides a rationale for the selection of a case-mix adjustment tool. Finally, this chapter provides a summary of the patient and state Medicaid factors that could potentially affect the quality of chronic care.

Quality Measurement

Quality improvement has received considerable attention in U.S. health policy initiatives. As healthcare costs continue to rise, efficiently managing healthcare service delivery becomes paramount for all stakeholders of healthcare, especially payers. This notion, coupled with the national focus on quality healthcare and empirical evidence indicating gaps in existing healthcare quality, provides an impetus for quality improvement. A crucial element of quality improvement is the availability of valid measures of quality of care. In the last decade, several researchers and organizations have developed valid indicators for measurement of quality.[5,10,11,14,15] Quality measurement and reporting at the levels of the patient, physician, payer, hospital and facility have received considerable attention.[1,2,3,6] Most quality indicators developed to be used at these levels are related to medication safety and effectiveness. These measures can be readily used to

assess performance of Medicaid programs on medication use-related metrics. Health plans are currently using technical quality indicators related to medications such as use of inappropriate medications in the elderly and adherence with chronic care medications to monitor quality at the plan level.[14] Centers for Medicare & Medicaid Services (CMS) reports the quality of medication use in their Part D plans. Yet, quality monitoring at the Medicaid level has received less attention.

Need for Quality Measurement

The U.S. healthcare system is consistently ranked first globally in terms of total healthcare spending. [16] However, the U.S. population has poorer health outcomes compared to many other developed nations.[16] As healthcare costs continue to grow, efficiently managing healthcare delivery becomes paramount for all healthcare stakeholders. Efficient management of healthcare entails the provision of good quality healthcare, while reducing costs. Quality improvement has been shown to reduce healthcare costs in the long-term, which explains the surge in healthcare quality improvement efforts in the US.[17] To improve quality, it is important to first define and measure quality. Once quality is measured, the information can be used to improve quality in the following ways: (1) healthcare organizations can use this data to identify opportunities for improvement, (2) healthcare organizations can also use the data to monitor the success of their quality improvement initiatives, (3) regulatory and accrediting bodies can use this data to retain accreditation, (4) provider payment can be adjusted based on the quality of care delivered, and (5) quality reports can be made publicly available, so that consumers can make informed choices.[17] Similar strategies can be applied to state Medicaid programs to stimulate quality improvement.

Traditionally, managed care plans and other payers relied on structure components for regulation purposes. An emphasis on process and outcomes measures has the potential to improve quality of chronic care greatly. Pay-for-performance and public reporting are two other ways to influence chronic care quality provided by payers.[18] These three strategies were successfully implemented in the context of physicians, hospitals, and long-term care facilities, demonstrating improvements in the quality of medical care provision.[19-21]

Definition and Measurement of Quality of Medical Care

A standard definition of healthcare quality is "the degree to which health services for individuals and populations increase the likelihood of desired health outcomes and are consistent with current professional knowledge".[22] While quality has been defined in numerous other ways, the main purpose of defining and measuring quality is the identification of opportunities for quality improvement.[17]

Donabedian's framework of quality is often used for healthcare quality measurement.[23] According to Donabedian, quality of care can be measured using three components: (1) structure, (2) process, and (3) outcome measures. Structural elements of medical care are necessary but not sufficient to ensure quality care. For example, structural components of care in the pharmacy setting would include computer systems for tracking patient information, patient counseling areas, and professional licensure. Process measures refer to the services and products provided by the pharmacist to the patient and include activities such as identifying and resolving drug-related problems, providing patient counseling and developing a caring relationship with the customer. Outcomes are perceived to be the ideal indicator of quality of medical care provided to patients. Donabedian defined outcomes as "a change in a patient's current and future health status that can be attributed to antecedent healthcare."[23] Outcomes can be broadly classified as

economic, clinical and humanistic outcomes.[24] Examples of outcome measures include cost of care, emergency room visits, mortality, and health-related quality of life among others.

Existing Quality Measures

In the last two decades, this stream of research has moved from justifying the need for quality measurement to the development of appropriate measures of quality. Quality is monitored using quality indicators or performance indicators. While there is a difference between the terms performance and quality, most researchers use quality and performance interchangeably. The main difference between the two is that quality indicators infer judgment about the care provided whereas performance indicators are devices for monitoring care provided without any inferences about quality.[25]

Quality measurement using administrative databases

Historically, numerous sources of data have been used to construct quality indicators. The most common data sources are medical records, administrative claims, and patient reports. Using medical record abstractions and patient reports to construct quality indicators is often labor intensive, slow and expensive to obtain. Administrative claims, though not originally intended for quality assessment purposes, provide an attractive option to measure quality.[26]

Administrative claims databases contain information on the patient, provider and facility and can be aggregated to measure quality at any of these levels. Claims data are routinely collected and are easily accessible due to their electronic format, but often remain underused in measuring quality of care. A major concern with this source of data is the lack of clinical information, which is necessary for calculating several existing quality indicators. Quality organizations are trying to develop quality/performance indicators that could be used solely with

administrative claims databases to improve the efficiency of quality measurement.[27,28] State Medicaid programs have convenient access to this data, and could readily use it for quality measurement purposes.

Technical Quality Indicators

Most existing technical quality measures are based on the processes, intermediate outcomes or final outcomes of care. Quality improvement efforts are intended to improve health outcomes. Therefore, outcomes should be the ideal measure of quality. However, the most commonly used technical quality measures are related to the process of care.[29,30] The advantage of process measures is that they are strongly influenced by provider behaviors rather than patient characteristics. Process measures that have an established link to health outcomes are very useful measures of quality of care.[7] Health outcome measures are not typically used in quality measurement for the following reasons: (1) important outcomes are sometimes rare, requiring large sample sizes to make statistically valid conclusions, (2) outcomes are affected mainly by patient factors and not providers, and (3) outcomes are observed after a long periods of time (e.g., ESRD in diabetic patients). Intermediate outcomes are observed immediately after clinical interventions, but they are often not available in administrative claims databases (e.g., blood pressure control, HbA1c levels). However, outcomes are preferred over process measures in cases where the outcome has a strong link to the process and occurs in a timely, efficient, and reliable manner.[7] Since this study focuses on chronic care quality where outcomes are observed after long periods of time, and because only two years of data will be used, process measures will be used to assess quality of chronic care.

Technical quality measures and performance measures are commonly expressed as ratios at the provider level. The denominator in the ratio is the number of patients eligible for the

measure and numerator is the number of patients who achieved a specified goal. For instance, if we were to construct an indicator for adherence with oral hypoglycemic medications at the state Medicaid level, the numerator would be the number of Medicaid beneficiaries adherent to oral hypoglycemics and the denominator would be all patients with diabetes using oral hypoglycemic enrolled in the state Medicaid program.

AHRQ Technical Quality Measures for Medicaid-Eligible Adults

In December 2010, AHRQ compiled a list of initial core set of health quality measures applicable to Medicaid-eligible adults for public comment as a part of section 2701 of the Affordable Care Act.[5] AHRQ created a subcommittee to enable compilation of new measures, and the following criteria were used to select measures: (1) the scientific acceptability of the measure, (2) feasibility of use by Medicaid and (3) importance to Medicaid programs.[5] Broadly, these measures fall in the areas of prevention and health promotion, management of acute conditions, management of chronic conditions, family experiences of care, and access. Several of the measures proposed by AHRQ for the management of chronic conditions are related to medication use, and they overlap with the quality indicators endorsed by the Pharmacy Quality Alliance (PQA) to assess community pharmacy quality.[10,11] AHRQ chronic care quality measures include three additional measures related to beta-blocker use post-myocardial infarction, statin use in coronary artery disease patients and persistence with antidepressant therapy. While the technical specifications have not been developed for these measures, this study will borrow the specifications provided by PQA and NCQA, where appropriate.

The Affordable Care Act mandates the Secretary of Health and Human Services (HHS) to establish a Medicaid quality measurement program by January 2012.[5] All state Medicaid programs will start reporting these measures in 2013, and this study will test the feasibility of

using Medicaid data to assess performance on these measures. This study will also provide baseline estimates of quality on management of chronic conditions, forming the foundation for assessing changes in quality scores in the future.

However, there is no existing evidence on the usability of Medicaid data to assess chronic care quality using the AHRQ proposed chronic care quality measures. As previously stated, data from 45 individual state Medicaid programs and the District of Columbia will be used in this study. Medicaid managed care penetration is variable across states, and some programs like Mississippi and North Carolina Medicaid were exclusively fee-for-service until 2011. The exhaustiveness of the data submitted by Medicaid managed care plans has not yet been studied, and the quality and quantity of Medicaid data across states available to researchers is variable. Therefore, the usability of Medicaid data in assessing chronic care quality may vary across states depending on the population size and managed care penetration.

Rationale for Measure Selection

It is important to assess the appropriateness of selected indicators before using them to assess quality. An ideal indicator should be clearly defined, quantitative, reliable, and tightly linked.[7,17] An indicator that is clearly defined provides definitions for all variables and explicates the inclusion and exclusion criterion for selection of cases for the denominator and the numerator.[7] Appropriate technical specifications are available for the AHRQ proposed measures that overlap with the existing PQA measures. For the three additional measures, a detailed methodology, including definitions, inclusion and exclusion criteria will be provided in this study. If an indicator is tightly linked, meaning the link between the process and outcome is clearly established, then the indicator becomes clinically meaningful and actionable. Indicator selection for this study was based on the aforementioned criteria. Additionally, an evidence of

sufficient variability in the quality indicator across Medicaid programs and a room for improvement in performance are required for justifying the measurement of quality indicators.[27] For instance, if existing research suggests that over 95% of the patients with coronary artery disease are using statins, it is not very meaningful to focus quality improvement efforts in that area. Though applicable to all quality indicators, the aforementioned criteria are particularly important while assessing chronic disease care quality indicators.

Therefore, areas reported to have substantial suboptimal quality should be the focus of future research. Numerous studies have indicated a strong link between the AHRQ chronic care process measures and health outcomes.[31-76] Existing literature provides sufficient evidence that there is a considerable room for improvement in the quality of medical care provided in these areas. The following section evaluates the AHRQ chronic care quality indicators on these criteria.

Persistence of beta-blocker treatment after a heart attack

Studies since the 1980s have reported that use of beta-blockers post myocardial infarction (MI) can improve survival and reduce the risk of secondary MI.[31-33] However, multiple studies have also demonstrated that beta-blockers are underused in post MI patients.[34,35] While there has been an increase in the prescribing of beta-blockers post-MI at hospital discharge from 63% in 1996 to 93% by 2002, the long-term outpatient use of beta-blockers is reported to be suboptimal.[36-38] Sustained used of beta-blockers has been shown to improve survival after MI in several randomized trials and observational studies. In a study examining the use of beta-blockers for one year after acute MI, only 45% of patients were found to be adherent to beta-blockers and the biggest drop in the percentage of patients adherent to beta-blockers was between 30 and 90 days.[36] Ackincigil et al.[39] found that only 50% of patients continued using

beta-blockers post MI. A recent study reported that adherence rates for beta-blockers have significantly improved over time but remain unsatisfactory.[40]

Statin use in people with coronary artery disease

The benefits of lipid lowering therapy using statins in patients with coronary artery disease are well documented.[41-44] Significant reductions in cardiovascular morbidity and mortality have been demonstrated in patients with and without coronary artery disease (CAD) using statins across large randomized clinical trials. For example, the West of Scotland Coronary Prevention Study reported a 31% risk reduction of non-fatal MI or death in patients with coronary artery disease.[41] The Scandinavian Simvastatin Survival Study (4S) demonstrated a 34% risk reduction in coronary events in statin-using patients with CAD.[42] Several other large cardiovascular outcome studies have shown significant reductions in death due to CAD with statin use.[43,44] Despite the compelling evidence of the benefits and clinical guidelines recommending statin therapy, there are wide gaps in statin use among patients with CAD.[45] In a survey conducted in nine European countries, only 32% of patients with coronary artery disease were found to be using a lipid-lowering drug.[46] Whincup et al.[47] reported that the prevalence of lipid lowering drug use was 29% among patients with a documented CAD. In a large study exploring treatment rates among 48,586 patients with CAD from 140 medical practices, only 39% received a lipid-lowering medication.[48] All these studies signify a considerable room for improvement.

Use of appropriate medication in people with asthma

The National Asthma Education and Prevention Program's Expert Panel Report 3 (EPR 3) on the diagnosis and management of asthma recommends the use of inhaled corticosteroids

(ICS) for the treatment of persistent asthma.[49] If asthma is not controlled by the use of ICS alone, long-acting β_2-adrenergic agonist (LABA) is added, but LABA should not be prescribed alone. Use of ICS is related to a decrease in asthma-related morbidity and mortality.[50] Still, patients are reported to be overusing reliever medications and under using preventer medications.[51] Kandane-Rathnayake et al.[52] explored guideline adherence for middle-aged adults with persistent asthma and found that only 29% of patients were prescribed ICS.

Persistence with antidepressant medications

Persistence with antidepressant therapy is crucial for consolidating treatment response and for reducing the risk of relapse. Clinical practice guidelines recommend that antidepressant medications should be used for at least four - nine months post remission of acute symptoms.[53-56] Suboptimal duration of antidepressant medication use is reported to increase the risk of relapse.[57,58] However, growing evidence indicates that majority of patients discontinue antidepressant medications during the first few weeks of treatment. For example, between 29 - 42% of patients prematurely discontinue medications after four weeks, and the discontinuation rate is between 63-76% after six months.[59-63] A high rate of discontinuation was also reported in Medicaid patients, with 70% patients discontinuing therapy in the first six months. Patients who discontinued antidepressant therapy had a 77% higher risk of relapse than patients who continued to use their medications in this study.[64]

Adherence with oral hypoglycemic, antihypertensive and antihyperlipidemic medications

Medication adherence is a major concern across all therapeutic categories with estimates of nonadherence ranging between 30 – 60%.[65] This finding was corroborated by another study reporting that the medication nonadherence averages 50% among chronic disease patients.[66]

Average adherence with prescribed therapy for type 2 diabetes patients was reported to be 67.5%.[65] A meta-analysis of studies exploring adherence to oral hypoglycemic medications reported that adherence ranges between 36-93%.[67] Similar trends of non-adherence were reported with lipid lowering therapy.[68] In another study, only 50% of patients were reported to be adherent to antihyperlipidemics six months after treatment initiation and the adherence rate dropped to 30-40% within one year.[69] In case of antihypertensive medications, the adherence rate is also suboptimal, ranging between 50-70%.[70]

Adherence to oral hypoglycemic medications has been linked to glycemic control and type 2 diabetes outcomes across several studies.[68,71] Schectman et al.[71] reported that for every 10% increase in drug adherence, HbA1c significantly decreased by 0.16%. Adherence to antihypertensive and antihyperlipidemic medications is linked to reduction in risk of coronary heart disease, stroke and death across numerous clinical trials and meta-analyses.[72-76]

Overall, there is sufficient variation in quality across patients for all the AHRQ measures included in this study. However, few studies explored if these variations in quality at the patient level translate to practice variations at the provider level or the payer level, or if it is mainly due to the patient case-mix. Also, there is a tight link between the AHRQ process measures and health outcomes, making these measures good indicators of quality.

Case-Mix Adjustment

It is important to report case-mix adjusted performance scores because Medicaid programs with a greater proportion of younger, sicker individuals, females or people from racial minorities are expected to perform poorly on these measures.[77-80] Risk adjustment methods will help explain variation in chronic care quality related to patient factors, so that the remaining

differences in quality measures can be attributed to other factors, such as the physician, pharmacist or the payer. Patient selection bias is another factor to adjust for when providing performance scores. Greenfield et al.[78] demonstrated that comparison of provider groups may be inaccurate if adjustment for patient case-mix and provider level clustering is not taken into account. Failure to control for payer level clustering by using techniques like multi-level regression analysis will cause standard errors of regression coefficients to be underestimated.[77]

Case-mix adjustment has been used to adjust performance of providers, facilities, or healthcare systems for making comparisons on the dimensions of access, efficiency and quality of care.[81] Several studies have found that adjusting for patient demographics, disease severity and comorbidity burden changes provider's performance profile.[78-80,81-83] Hofer et al.[6] demonstrated that physicians can improve their performance profile by preferential patient selection. A recent study found that adjusting for patient characteristics and treatment opportunities improved hospital rankings on indicators assessing adherence to treatment guidelines for acute MI.[80] Importance of case-mix adjustment was also demonstrated in evaluating health plan performance on chronic care provided to Medicaid enrollees.[82] Similarly, application of risk-adjustment methodologies to adjust for differences in patient's case-mix will be important in assessing the quality provided by state Medicaid programs.

A plethora of measures are available to adjust for patient case-mix in the outpatient care setting. Broadly, these measures can be classified into three components: (1) health status, (2) patient health behaviors and psychosocial factors and (3) contextual factors.[79] Sociodemographic factors (e.g., age, gender, insurance type) and contextual factors (e.g., number of specialists available in a given area, median income in the county) are commonly used dimensions of risk in studies using administrative claims databases, because this information is readily available.[83]

Ideally, a comprehensive understanding of population's health status using clinical assessment and a complete sociodemographic profile would provide an accurate assessment of risk. However, such methods are often cost prohibitive, and several alternative tools for risk assessment that use administrative claims are available to researchers.

Diagnosis-based or pharmacy-based measures are widely accepted to classify patients into different risk categories.[84-87] Adjusted clinical groups (ACG) and diagnostic cost groups/hierarchical condition categories (HCC) are the commonly used diagnosis-based measures.[86] ACG classifies patients into 32 clinical groups called Ambulatory Diagnostic Codes (ADGs) based on disease duration and severity. Each person is further classified into mutually exclusive categories based on age, gender and total number of ADGs. DCG model assigns patients to Condition Categories (CCs), based on disease type grouped by cost and clinical relation.[81,88] This model also includes several disease interaction terms and child-specific conditions are weighted. However, the completeness and reliability of ICD-9 codes could vary across providers, facilities and healthcare systems. Other measures like Charlson's comorbidity index (CCI), and Elixhauser's comorbidity index (EI) are more commonly used to adjust for comorbidity burden.[68,86] The applicability of diagnosis-based risk-adjusters like ACG and HHC is mostly seen while modeling healthcare utilization and costs.[86]

Alternate approaches to case-mix adjustment use pharmacy claims data. The most extensively used pharmacy-based case-mix adjustment measure is RxRisk.[87] It is also used as a disease severity measure.[81] Patients are classified into 42 non-mutually exclusive disease conditions based on the medications used to treat chronic conditions.[86] Other measures like chronic disease score, pharmacy cost groups, and pediatric chronic disease score have been used to adjust for case-mix and disease severity.[89-91] In its earlier applications, the RxRisk model was

predominantly used as a risk-adjustment tool when modeling cost and utilization data.[86] More recently, the RxRisk measure was found to be a better predictor of adherence to diabetes medications compared to CCI, EI, and Health Related Quality of Life comorbidity index.[92,93] Chronic disease score and CCI index were mostly used in predicting medication use across several studies.[68,94]

The RxRisk instrument coupled with demographics (age, sex, race) will be used to adjust for patient case-mix in this study. Pharmacy-based models are useful especially for populations with drug benefits like Medicaid. One main advantage of using pharmacy claims based risk adjuster is that stable chronic diseases are better captured using drug records than by using ICD-9 codes. Patients with a stable chronic condition may not visit a physician because most health plans do not require a physician visit to fill or refill a prescription. The choice of pharmacy-based risk adjustment is also driven by the fact that this study will predominantly use pharmacy claims data (except for two measures) so that payers can use this tool to create quality scores related to medication use for most of the measures being studied, using limited data.

Chronic Care Quality

An estimated 144 million people in the United States had one or more chronic conditions in 2010. Approximately 20% of these individuals have multiple chronic conditions.[95] The prevalence of chronic conditions is projected to increase in the near future. Cost of managing these conditions is estimated to be $1.8 trillion annually, which forms 75% of all healthcare spending.[96] As the cost of chronic conditions continues to rise, it is important to ensure that the quality of care can be improved and expenditures controlled. Most chronic diseases such as diabetes, cardiovascular disease, asthma, and depression can be effectively managed using lifestyle changes and medications, thereby improving health outcomes and reducing overall

expenditures.[68,97] For example, appropriate management of diabetes using oral hypoglycemic agents was shown to improve health outcomes and decreases expenditures.[68] On the other hand, a lack of adherence to asthma medications has been implicated in costly procedures and hospitalizations.[97]

Payers play an important role in the management of chronic conditions by providing better access to pharmaceutical care, disease management, and medication therapy management. Therefore, it is appropriate to measure quality of chronic care at the payer level. If there is an evidence of variation in chronic care quality provided by the various state Medicaid programs, it reflects a scope for quality improvement. There is also a paucity of studies examining variation in chronic care quality across state Medicaid programs. Evaluating variation across state Medicaid programs can provide insights into healthcare quality issues relative to resource allocation and access to care.

Variation in quality of care as measured by medication adherence or persistence and use of inappropriate medications provided to patients with chronic conditions across provider, facility and health system level have been reported.[9,98,99] For instance, multilevel modeling studies assessing variation in diabetes care contributed by the patient, provider and facility have found that a large proportion of variation was explained by patients, with less but substantial variation explained by clinician and facility.[100,101] Most studies using multi-level data only report the amount of variation in quality that is attributable to each of the levels of analysis, typically the patient, provider and facility.[102] A wide range of studies also assessed patient characteristics associated with quality of care, including chronic care.[68,97] However, characteristics of the payer associated with suboptimal quality are rarely reported.

Factors affecting quality of care

There is a paucity of studies investigating the provider, facility and payer/program characteristics as determinants of variation in quality of care, in general. Fewer studies explored factors associated with chronic care quality using measures related to medication use. Factors like physician age and specialty at the physician level, and number of providers at the state level were some of the factors assessed in previous studies.[98-101] The conceptual framework for understanding the factors explaining chronic care quality is provided in Figure A-1 (APPENDIX). The choice of patient and state characteristics presented in the model are based on prior research in this area and data availability.

Patient Characteristics

The Aday-Andersen model for healthcare utilization is widely used to understand patient characteristics explaining a particular medication behavior, like medication adherence.[103] This model categorizes the determinants of healthcare utilization behavior into three components: predisposing, enabling, and need-related factors (APPENDIX - Figure A-2). Predisposing factors such as age, gender or race are indicative of a person's propensity to utilize healthcare. Enabling factors are related to the person's ability to gain access to services, such as income, insurance, and other access variables. Need-related factors are defined as the disease severity or comorbidity burden. Since, this study population is only comprised of Medicaid beneficiaries, people enrolled in a state Medicaid program may not have meaningful differences in enabling factors like access and income. Few access variables (e.g., proximity to the pharmacy/physician, rural or urban residence) may still be useful in understanding medication use in the Medicaid population. However, only predisposing (age, sex, and race) and the need-related (RxRisk score) patient characteristics will be included in this study to adjust for patient case-mix.

State Medicaid Characteristics

State Medicaid policies affect beneficiary's access to and use of healthcare services, so there are likely to be differences in chronic care quality across states. In a study examining if adult healthcare access and use varied across state Medicaid programs, case load per state, physician reimbursement, co-payments, and limits on number of physician visits or hospital days were some of the factors found to affect use.[13] Tang et al.[104] also used 50 state Medicaid data to assess state level factors (such as, state managed care penetration, primary care and mental health providers and Medicaid reimbursement rate) related to unmet mental healthcare need for Medicaid children and reported that none of the state-level variables were significant in predicting unmet need. Additionally, physician supply is a factor to consider when assessing quality of care provided to Medicaid patients because fewer physicians accept Medicaid patients relative to Medicare and privately insured patients, mainly because of low reimbursement rates.[12]

Since chronic care quality often depends on access to care, primary care physician supply, specialist supply, pharmacy benefits, presence of medication therapy management programs, and managed care penetration are some of the factors that may affect quality measures being used in the study.

CHAPTER III – RESEARCH METHODOLOGY

Data Source

Medicaid is a federal aided, state-operated program providing healthcare coverage for certain indigent or low-income individuals and families. Under broad federal guidelines, states devise their own programs by establishing eligibility standards, determining the scope of services provided, setting payment rates and administering their own programs.[105] Therefore, each state's Medicaid program is unique. Medicaid data is made available for research purposes through the Research Data Assistance Center (ResDAC).[106] States submit provider claims and payments electronically to the Centers for Medicaid & Medicare Services (CMS) through Medicaid Management Information Systems (MSIS). ResDAC is a CMS contractor that provides de-identified data in the form of Medicaid Analytic Extract (MAX) files which are extracted from the MSIS system. The MAX files contain beneficiary level enrollment, utilization and expenditure data. Medicaid MAX data files from 45 states (except Alaska, Montana, North Dakota, South Dakota, and Wyoming) with over 100,000 enrollees and the District of Columbia for the years 2006 and 2007 were requested from ResDAC for the purpose of this project. Medicaid eligibility, pharmacy claims, medical claims and inpatient claims files were used in this study. The components of each of the files are enlisted below:

1. Person summary file: Person-level file with demographic information and eligibility periods for the beneficiaries.

2. Pharmacy claims file: Event level file with information on prescription dispensed, amount paid, quantity dispensed, prescription date, and prescribing physician.

3. Other services file: Event level file with information on the outpatient facility visits (outpatient hospital, physician office, ER), diagnoses, amount paid, and provider type.

4. Inpatient services file: Event level file with information on the admission and discharge dates, diagnoses, and amount paid.

Study Design

This study was a retrospective analysis of administrative claims data from Medicaid programs for 45 states and the District of Columbia for the years 2006 and 2007. Eleven measures from the AHRQ initial core set related to management of chronic conditions using medications were studied, including – (1) persistence with beta-blocker medications post-MI (MI), (2) statin use in people with CAD (CAD), (3) use of ICS or similar medications in individuals with persistent asthma (ICS), (4) persistence with antidepressant medications, acute and chronic (AD), (5) adherence to ACEI/ARBs (ACEI), (6) adherence to beta-blockers (BB), (7) adherence to calcium-channel blockers (CCB), (8) adherence to statins (STAT), (9) adherence to biguanides (BIGU), (10) adherence to sulfonylureas (SU), and (11) adherence to thiazolidinediones (TZD). An additional measure not included in the AHRQ core set, assessing beta-blocker use in people with MI (MI1) was also computed. Medicaid data for the year 2006 were used for computing the MI and AD measure scores. Detailed definitions of the 11 selected measures are provided in Table A-1 (APPENDIX). Medicaid beneficiaries for whom at least one of the measures can be assessed were identified using a combination of drug claims data and diagnosis codes using the patient selection criteria (see Eligibility Criteria).

Eligibility Criteria

Medicaid beneficiaries between 18-65 years of age by the end of 2007 and with continuous enrollment in the particular state Medicaid program during the measurement period were included if they met the indicator specific eligibility criteria. To be considered continuously enrolled, a Medicaid beneficiary should not have a gap of more than one month during each study year. Additionally, dual-eligible patients that are enrolled simultaneously in Medicaid and Medicare were excluded from this study. A combination of National Drug Codes (NDCs) and International Classification of Disease, 9th Revision, Clinical Modification (ICD-9- CM) codes were used to identify beneficiaries eligible for each measure. NDCs for the list of medications in Table A-2 were compiled from Multum drug database and matched to drug claims in the pharmacy claims file to flag medication use. Similarly, ICD-9-CM codes were matched to diagnosis claims in the inpatient and medical claims files to identify cases for MI and CAD measures. Additional criteria for patients to be eligible for each measure are listed below:

- MI – A primary or non-primary diagnosis code for myocardial infarction (ICD-9 code: 410;412) in the inpatient claims file between April 1, 2006 and March 31, 2007. Since, beta-blockers have multiple indications, persistence with beta-blockers were observed post identification of patients with MI. Patients filling a prescription for a beta-blocker within one month after discharge from the hospital were included in the denominator.
- CAD – A primary or non-primary diagnosis code for coronary artery disease (ICD-9 codes: 410.X0, 410.X1, 410.X2 where X=0-9; 412;413;413.10;413.90;414;414.0X where X=1-7;414.80;414.9;v45.81;v454.82) in 2007.
- ICS – At least two prescriptions for asthma medications (Asthma – A, included in Table A-2 (APPENDIX)) within four months of one another. Patients who filled at least one

prescription for a COPD medication, pulmozyme or a nasal steroid medication (Asthma – B, included in Table A-2) were excluded from this measure.

- AD – At least two prescriptions for an antidepressant medication (included in Table A-2) filled on two unique service dates, with the first prescription filled between May 1, 2006 and April 30, 2007 and no prior use of antidepressants for at least three months.
- ACEI – At least two prescriptions for ACEI/ARB or ACEI/ARB combination medications (included in Table A-2) filled on two unique dates of service at any time in 2007.
- BB – At least two prescriptions for beta-blocker or beta-blocker combination medications (included in Table A-2) filled on two unique dates of service at any time in 2007.
- CCB – At least two prescriptions for calcium channel blocker or calcium channel blocker combination medications (included in Table A-2) filled on two unique dates of service at any time in 2007.
- STAT – At least two prescriptions for statin or statin combination medications (included in Table A-2) filled on two unique dates of service at any time in 2007.
- BIGU – At least two prescriptions for biguanide or biguanide combination medications (included in Table A-2) filled on two unique dates of service at any time in 2007.
- SU – At least two prescriptions for sulfonylurea or sulfonylurea combination medications (included in Table A-2) filled on two unique dates of service at any time in 2007.
- TZD – At least two prescriptions for thiazolidinedione or thiazolidinedione combination medications (included in Table A-2) filled on two unique dates of service at any time in 2007.

Medications dispensed in key therapeutic classes were used instead of diagnosis claims for identifying eligible patients for all except two measures, MI and CAD. There are two reasons for this approach: (1) a person with a stable chronic disease is better captured from their drug records than using ICD-9 codes with only one year of data. Patients with a stable chronic condition may not visit a physician because most health plans do not require a physician visit to fill or refill a prescription. (2) We used pharmacy claims only for those measures wherein confirming the diagnosis is not germane to questions about the validity of the measures.[27] The majority of these measures were related to persistence/adherence with medications and did not require a diagnosis, except MI and CAD. For instance, MI measures the persistence of beta-blockers post myocardial infarction, but beta-blockers are used for managing multiple cardiovascular conditions. Patient post an acute MI might be more persistent with their beta-blocker therapy compared to those with mild hypertension. Therefore, we could be overestimating or underestimating the performance measure without the inclusion of a MI diagnosis. Also, if it can be shown that a majority of these indicators can be constructed using pharmacy claims alone, Medicaid can request community pharmacies to report on these measures.

Patients with non-acute stays anytime during the measurement year were excluded from the analysis. Patients with at least one claim in the long term care services file were excluded. Non-acute stays were identified from the other services file using UB revenue codes related to hospice, skilled nursing facility, rehabilitation, respite, residential substance abuse treatment facility and psychiatric residential treatment center.

A list of UB revenue codes is provided in Table A-3 (APPENDIX). Patient in non-acute care facilities listed above are a select population that is sicker and may demonstrate different medication use behavior as they are under constant supervision.

Study Variables

Measures

The numerator and denominators for the 11 selected measured are provided in Table A-1 (APPENDIX). Medication adherence (ACEI, BB, CCB, LLD, BU, SU, TZD) were measured as proportion of days covered (PDC). PDC is one of the most common methods of adherence measurement with increasing use in literature.[107] PDC is calculated as the total number of days with medications on hand divided by the specified time interval. The denominator for the PDC measure is the patient's measurement period, defined as the index prescription date to the end of the calendar year. PDC is multiplied by 100 to yield a percentage and ranges between 0 and 100%. Typically, a PDC measure is adjusted for hospitalizations by subtracting the number of days spent in a hospital from the numerator and the denominator. However, since use of non-pharmacy data is limited in this study, hospitalization stay adjusted PDC was not computed.

PDC were measured for all seven therapeutic classes for each eligible patient. Switching between medications within the same therapeutic class during the measurement period were treated as continuous use of that class of medication, to reduce the complexity of adherence measurement. Concomitant use of medications from different therapeutics classes and switching across therapeutics classes is common in individuals using oral hypoglycemic and antihypertensive medications. Therefore, PDC measure scores for oral hypoglycemic and antihypertensive medication classes were reported. PDC less than 80% was considered

nonadherent, a common threshold used in adherence research for the therapeutic classes in question.[108] Implicit to this approach is an assumption that patients who initiate a particular class of medication will maintain the therapy throughout the measurement period. A limitation to this approach is that it is difficult with claims data to determine whether or not a patient was advised by their physician to discontinue or switch therapy.

Persistence represents the time over which a prescription is filled by the patient. For persistence measures (MI and AD), treatment was considered continuous if the total gap without medications was less than 51 days over a period 6 months. Persistence with medications was measured as a dichotomous variable, with a total gap in medication therapy of 51 or more days (AD) during the 6 month measurement period considered non-persistent. A limitation of using claims data to capture persistence is that it is not possible to determine if the medication was discontinued as per physician recommendation. An additional measure related to effectiveness of antidepressant therapy during an acute phase was computed. For the AD acute measure, total gap in medication therapy of 30 or more days (AD) during a 12 week measurement period was considered non-persistent.

For the ICS measure, patients were included in the numerator if they have at least one prescription for inhaled corticosteroids, nedocromil, cromolyn sodium, leukotriene modifiers or methylxanthines during the measurement year. Patients were included in the numerator for the CAD measure if they had at least one prescription for a statin during the measurement year. Additionally, a summary composite measure of chronic care, defined as the total number of measures for which the individual received good care (i.e., adherent, persistent or used the recommended therapy) divided by the total number of measures the subject is eligible for were calculated for all Medicaid adults meeting the eligibility criteria.

State Medicaid performance on each measure were calculated as the number of patients who received good quality care (i.e., adherent, persistent, received or used the recommended therapy) divided by the number of patients eligible for the measure. A state-level measure was also created for oral hypoglycemic and antihypertensive medications. Additionally, the composite score developed at the patient level as a summary measure of chronic care was aggregated to the state Medicaid level.

Patient Characteristics

Information on patient demographic variables was obtained from the personal summary file.

Age: Age was calculated as the difference between the beneficiary's date of birth field 'DOB07' and the end of the measurement year '31st December 2007'. It was treated as a continuous variable.

Gender: Male or female. The field 'SEX07' was utilized to identify patient's gender.

Race/ethnicity: Four major racial and ethnic groups were considered: Non-Hispanic Whites, African Americans, Hispanics, and Other races. Based on the racial distribution of the population across the eligible states, we collapsed groups with a small representation into "other races". The field 'RACE107' was used to identify race/ethnicity.

RxRisk: RxRisk was used as the comorbidity burden measure. RxRisk, which is an extension of the chronic disease score, uses pharmacy claims to quantify patient's comorbidity burden. RxRisk system includes 42 chronic disease categories related to adults. A list of the chronic disease categories related to adults and all the medication classes that belong to each category is provided in Table A-4 (APPENDIX).

To construct the RxRisk score, indicators were created for the different chronic disease categories. Each individual is flagged as having a chronic condition by matching the NDC codes (for medications related to the chronic disease categories) to their prescription records anytime during the measurement period. These indicators were summed to create the unweighted RxRisk score for the individual. Additionally, a weighted RxRisk score was created using empirical weights based on the predictive validity of the chronic conditions for mortality obtained from Johnson et al.[109]

Data Management

MAX data files are fully de-identified with encrypted beneficiary identification numbers. MAX data files were made available to the researcher in SAS format. Beneficiaries without a single claim in the study period and beneficiaries with missing prescription service date or product service date in pharmacy claims were excluded. The data were checked for duplicate claims and only the first record was retained in the case of duplication. Data management was conducted using SAS software version 9.2.[110] Research and analytical procedures were approved by the University of Mississippi Institutional Review Board.

Statistical Analysis

Objective 1: To assess the feasibility of using state Medicaid data to provide national estimates of chronic care quality on medication use-related metrics.

Each state is required to have a sample size of 30 or more patients in the denominator to be considered for a measure.[27] Performance rates on each measure were computed only for those states meeting this criterion. Additionally, at least 10 states should meet the minimum sample size criterion for reporting the performance rate on each measure at the national level.

Objective 2: To estimate crude and case-mix adjusted scores of chronic care quality on medication use-related indicators for state Medicaid programs.

For the 11 measures, the crude performance score at the state level was calculated as the number of patients in the numerator divided by the number of patients in the denominator for the particular measure. Additionally, a summary composite measure of chronic care (expressed as a proportion), defined as the total number of care opportunities in which a beneficiary received good care (i.e., adherent, persistent or used the recommended therapy) divided by the total number of care opportunities in the state. This measure provides a summary of the proportion of care opportunities fulfilled in each state.

The study uses statistical risk-adjustment, also known as case-mix adjustment, to control effects of confounding variables seen in patients. A hierarchical multivariable logistic regression analysis was performed to adjust performance scores (on the composite measure and the 11 AHRQ process measures) for patient case-mix, similar to the method proposed by Mehta et al.[80] Multilevel models are appropriate for this data, considering the hierarchical structuring of the data, where patients are clustered within states. Performance on each measure was adjusted for patient case-mix by including patient characteristics in the model. These characteristics include age, sex, race/ethnicity, and RxRisk score. The hierarchical approach treats states as random effects and allows adjustment for within-state correlation in the process measures.

Adjusted scores reflect the predicted mean scores for each state if they all had the same case-mix. For computing adjusted scores for the composite measure, opportunity-based data was used, meaning each measure for which patient is eligible will contribute an observation to the dataset. For instance, if a patient is eligible for the CAD, BB, MI measures and is eligible for the numerator (referred to as adherent hereafter) for BB and MI, the patient then has three

observations in the dataset, with two positive events. First, state-specific estimates of observed adherence rates were calculated as the mean of the predicted probability of adherence from the hierarchical model across all patients attributed to the state. Second, the expected adherence score in a given state was calculated as the mean of the predicted probabilities of adherence from the model with patient characteristics, without incorporating state as a random effect. To compute the adjusted performance score on the measure, state-specific observed adherence rate was multiplied by the overall observed adherence rate and divided by the state's estimated expected adherence rate.[80] Adjusted scores were computed similarly for the 11 measures, except each patient had one observation in the dataset.

To assess the difference between the adjusted and unadjusted scores, states were grouped into three categories: (1) low quality (if unadjusted score is less than the average unadjusted score and 95% CI of the unadjusted score does not contain the average unadjusted score; if case-mix adjusted score is less than 1 and the 95% CI of the case-mix adjusted score does not contain 1), (2) medium quality (if the 95% CI of the unadjusted score contains the average unadjusted score; if the 95% CI of the case-mix adjusted score contains 1), and (3) high quality (if unadjusted score is higher than the average unadjusted score and 95% CI of the unadjusted score does not contain the average unadjusted score; if case-mix adjusted score is higher than 1 and the 95% CI of the case-mix adjusted score does not contain 1). The agreement in classification based on unadjusted and adjusted performance scores was evaluated using Cohen's Kappa (κ) coefficient.

Additionally, the agreement in classification across the various measures was also evaluated using Cohen's Kappa (κ) coefficient by classifying states into low (bottom 20%), medium (middle 60%), and high (top 20%) performing states on the 11 measures. If the percentage of states whose classification would be changed by case-mix adjustment is substantial, then adjusted scores were used to describe the distribution of states for *Objective 3*.

Objective 3: To describe the distribution of Medicaid performance scores of chronic care quality on medication use-related measures.

Univariate statistics and coefficient of variation were used to describe the distribution of the performance scores across states. Median, mean, 10^{th} and 90^{th} percentiles were reported for each measure individually and for the composite measure. The median score is representative of the room for improvement in performance scores and the 10^{th} and 90^{th} percentiles identify the spread of scores across the states. The 90^{th} percentile represents the top performing Medicaid programs in the nation. The coefficient of variation is a frequently computed statistic in studies of practice variation and is calculated as the standard deviation divided by the mean and multiplied by 100, so that the result can be expressed as a percentage. Coefficient of variation provides an indication of the spread of data and is sensitive to outliers. Additionally, to determine if performance on one indicator influenced the performance on another indicator, states were ranked independently on all 11 measures and, Kendall's τ rank order correlations was estimated among indicators using state as the unit of analysis.

Objective 4: To illustrate cross-state variation in performance scores of chronic care quality on medication use-related measures.

Variation across states was described by generating state-level US maps to depict cross-state variations in performance scores on the 11 measures and the composite measure.

Objective 5: To assess the variation in chronic care quality that may be attributed to the patient and state Medicaid levels.

For the 11 measures, hierarchical logistic regression models were used to assess the variation in chronic care quality attributable to each level. The composite measure was fitted using a three-level hierarchical logistic regression models. Hierarchical models were used to account for the clustering of patients within states. Using ordinary least squares with nested data can yield biased parameter estimates and inefficient standard errors. Multilevel models, on the other hand, allow errors to be dependent within contexts.[43]

An intercept-only model, without any covariates, was constructed for each measure and Medicaid identifiers were entered for each patient as a random effect. ICC was estimated for the state level based on the formula given by Snijders and Bosker (1999)[112]. The coefficients reflect the proportion of the total variance in the dependent variable that is attributable to the state level.[43] In addition, residual ICC was computed from the models with patient level variables included in the model and state variable entered as a random effect. The multilevel analysis was performed using the PROC GLIMMIX procedure in SAS version 9.3.[110]

The standard cut-off point of a p-value of 0.05 was used in all analyses.

CHAPTER IV - RESULTS

Feasibility of using Medicaid data to provide national benchmarks of medication use-related quality measures

Medicaid data was obtained for 45 states and the District of Columbia for the years 2006 and 2007. States with less than 100,000 Medicaid beneficiaries including Alaska, Montana, North Dakota, South Dakota and Wyoming were excluded from this study. All states met the eligibility criteria of having at least 30 eligible patients in the denominator for all measures except MI1 and MI. Seven and thirteen states had less than 30 beneficiaries eligible for the MI1 and MI measures respectively. Also, most measures could not be generated for Michigan and Ohio due to incompleteness of the day supply field. In addition, AD, CAD and MI measures could not be computed for Maine because of missing inpatient hospital and other services claims. Inpatient and other services claims are not required for computing PDC and ICS measures, but are used in identifying patients with non-acute stays.

Overall, MI measure could be generated for 30 states, MI1 measure for 36 states, AD measures could be computed for 43 states, PDC for 44 states, CAD for 45 states and the ICS measure could be calculated for 46 states. All measures could be estimated for at least 10 states, thereby meeting the minimum sample size criterion for reporting the performance rate on each

measure at the national level. A summary of the usability of data obtained from ResDAC in computing medication use-related quality measures is presented in Table 1.

Table 1: Data Quality Summary

State\Measure	Measure Eligibility					Reason
	PDC	AD	CAD	ICS	MI	
# eligible states	44	43	45	46	30	
Alabama	✓	✓	✓	✓	✓	
Alaska	-	-	-	-	-	Data not requested
Arizona	✓	✓	✓	✓	✓	
Arkansas	✓	✓	✓	✓	✓	
California	✓	✓	✓	✓	✓	
Colorado	✓	✓	✓	✓	✓	
Connecticut	✓	✓	✓	✓	✓	
Delaware	✓	✓	✓	✓	×	
District of Columbia	✓	✓	✓	✓	×	
Florida	✓	✓	✓	✓	✓	
Georgia	✓	✓	✓	✓	✓	
Hawaii	✓	✓	✓	✓	✓	
Idaho	✓	✓	✓	✓	×	
Illinois	✓	✓	✓	✓	✓	
Indiana	✓	✓	✓	✓	✓	
Iowa	✓	✓	✓	✓	×	
Kansas	✓	✓	✓	✓	×	
Kentucky	✓	✓	✓	✓	✓	
Louisiana	✓	✓	✓	✓	✓	
Maine	✓	×	×	✓	×	40.7% missing claims (IP and OT files unavailable)
Maryland	✓	✓	✓	✓	✓	
Massachusetts	✓	✓	✓	✓	✓	
Michigan	×	×	✓	✓	×	36% claims with missing day supply field
Minnesota	✓	✓	✓	✓	✓	
Mississippi	✓	✓	✓	✓	✓	
Missouri	✓	✓	✓	✓	✓	
Montana	-	-	-	-	-	Data not requested

Table 1: Data Quality Summary (continued)

State\Measure	Measure Eligibility					Reason
	PDC	AD	CAD	ICS	MI	
# eligible states	44	43	45	46	30	
Nebraska	✓	✓	✓	✓	×	
Nevada	✓	✓	✓	✓	×	
New Hampshire	✓	✓	✓	✓	×	
New Jersey	✓	✓	✓	✓	✓	
New Mexico	✓	✓	✓	✓	×	
New York	✓	✓	✓	✓	✓	
North Carolina	✓	✓	✓	✓	✓	
North Dakota	-	-	-	-	-	Data not requested
Ohio	×	×	✓	✓	×	100% claims with missing day supply field
Oklahoma	✓	✓	✓	✓	✓	
Oregon	✓	✓	✓	✓	×	
Pennsylvania	✓	✓	✓	✓	✓	
Rhode Island	✓	✓	✓	✓	×	
South Carolina	✓	✓	✓	✓	✓	
South Dakota	-	-	-	-	-	Data not requested
Tennessee	✓	✓	✓	✓	✓	
Texas	✓	✓	✓	✓	✓	
Utah	✓	✓	✓	✓	×	26.5% with no claims*
Vermont	✓	✓	✓	✓	×	
Virginia	✓	✓	✓	✓	✓	
Washington	✓	✓	✓	✓	✓	
West Virginia	✓	✓	✓	✓	✓	
Wisconsin	✓	✓	✓	✓	✓	
Wyoming	-	-	-	-	-	Data not requested

*Data anomaly obtained from annual Medicaid Analytic Extract Claims Anomaly Tables generated by Mathematica Policy Research.

Overall, 1,538,448 Medicaid beneficiaries were eligible for at least one of the study measures. Demographic characteristics and co-morbidity indices of this patient pool are described in Table 2. The average age of Medicaid beneficiaries included in the study was 46.2

years (SD: 11.9) and 68.0% were female. A total of 45.1% of the study population was white, 25.0% black, and 16.8% Hispanic. Co-morbidities were assessed using weighted and unweighted RxRisk and Charlson's comorbidity indices.

Table 2: Baseline characteristics of the study population

Baseline Characteristics	**Prevalence(%) or Mean ± SD**
N	1,538,448
Age	46.19 (11.94)
Sex	
• Female	67.99%
• Male	32.01%
Race	
• White	45.06%
• Black	25.04%
• Hispanic	16.76%
• Other	13.14%
RxRisk	5.96 (3.01)
RxRisk (Weighted)	7.88 (5.13)
Charlson's Comorbidity Index (Unweighted)	1.00 (1.14)
Charlson's Comorbidity Index	1.57 (1.87)

The number of patients eligible for the 11 measures varied considerably from 4,140 patients in the denominator for the MI measure compared to 706,849 patients for the ACEI/ARB PDC measure. Table 3 provides a detailed overview of the number of patients in the numerator and denominator for all study measures.

Table 3: Overview of patient eligibility by measure type

Measure Description	2007	
	# patients in the denominator	% patients in the numerator
PDC (Proportion of Medicaid patients 18-65 years who met PDC threshold of 80 percent)		
ACEI/ARBs	706,849	52.34
Betablockers	458,852	49.54
Calcium Channel Blockers	324,536	50.94
Biguanides	317,670	44.23
Sulfonylureas	200,182	46.56
Thiazolidinediones	154,969	41.76
Statins	557,765	48.32
Antihypertensives	1,025,359	54.85
Oralhypoglycemics	410,126	50.16
AD Acute (Proportion of Medicaid patients 18-65 years newly diagnosed with MDD and prescribed antidepressant therapy persistent with the therapy for at least 3 months)		
AD-Acute	93,960	52.88
AD Chronic (Proportion of Medicaid patients 18-65 years newly diagnosed with MDD and prescribed antidepressant therapy persistent with the therapy for at least 6 months)		
AD-Chronic	93,960	31.28
CAD (Proportion of Medicaid patients 18-65 years diagnosed with CAD that received at least one prescription for a statin medication)		
CAD	145,391	62.50
ICS (Proportion of Medicaid patient 18-65 years with persistent asthma that received at least one prescription for an inhaled corticosteroid or similar medication)		
ICS	262,043	68.12
MI1 (Proportion of Medicaid patients 18-65 years with MI that filled a beta-blocker prescription within 30 days of discharge)		
MI1	8,293	49.13
MI (Proportion of Medicaid patients 18-65 years with MI that filled a beta-blocker prescription within 30 days of discharge persistent with therapy for at least 6 months)		
MI	3,945	59.15

Crude and case-mix adjusted scores on medication use-related quality measures

Angiotensin-Converting Enzyme Inhibitor/Angiotensin-Receptor Blocker (ACEI/ARB)

The PDC measure for ACEI/ARBs could be computed for 42 states and the District of Columbia. The number of patients eligible for the measure in each state and the proportion of patients meeting the 80% adherence threshold are shown in Table 4. The number of patients in the denominator ranged from 1,109 for New Hampshire to 136,018 for New York. The average adherence rate for the ACEI/ARB measure was 52.3% across all states, ranging from 42.1% of Medicaid patients being adherent in Mississippi to 69.0% in Vermont. States with scores above the national benchmark are highlighted in bold.

Table 4: Patients meeting the ACEI/ARB measure criteria by state

State	ACEI/ARB	
	#patients in the denominator	**%patients in the numerator**
Alabama	16,818	**52.52**
Alaska	-	-
Arizona	18,680	46.93
Arkansas	5,931	**52.72**
California	102,520	47.91
Colorado	4,143	45.84
Connecticut	4,998	**61.42**
Delaware	3,620	51.35
District of Columbia	2,814	49.08
Florida	23,544	**55.16**
Georgia	21,082	45.43
Hawaii	2,679	48.82
Idaho	1,457	**68.36**
Illinois	30,267	**52.35**
Indiana	9,463	50.27
Iowa	4,019	**60.74**
Kansas	3,392	**58.70**

Table 4: Patients meeting the ACEI/ARB measure criteria by state (continued)

State	ACEI/ARB	
	#patients in the denominator	%patients in the numerator
Kentucky	22,459	**61.05**
Louisiana	18,748	49.79
Maine	-	-
Maryland	13,333	49.00
Massachusetts	17,058	**54.24**
Michigan	-	-
Minnesota	6,671	51.46
Mississippi	12,254	42.10
Missouri	15,198	**57.97**
Montana	-	-
Nebraska	1,843	**58.87**
Nevada	2,085	**58.13**
New Hampshire	1,109	**63.21**
New Jersey	16,062	52.09
New Mexico	5,636	**53.00**
New York	13,6018	**53.40**
North Carolina	26,544	52.20
North Dakota	-	-
Ohio	-	-
Oklahoma	7,106	50.86
Oregon	1,541	**62.88**
Pennsylvania	8,960	**64.53**
Rhode Island	3,608	**57.59**
South Carolina	13,217	48.38
South Dakota	-	-
Tennessee	30,672	47.38
Texas	36,830	**53.73**
Utah	2,677	**53.16**
Vermont	2,410	**68.96**
Virginia	12,682	**58.07**
Washington	13,011	**59.93**
West Virginia	11,090	**60.72**
Wisconsin	12,600	**53.77**
Wyoming	-	-
Total	706,849	52.34

Demographic characteristics and co-morbidity measures for the eligible patient population are described in Table 5. Approximately 64% of the patients were women, and the mean age (± SD) was 50 (± 10) years. Non-adherent patients were slightly younger and a greater proportion of non-adherent patients were Black and Hispanic compared to the adherent patients. The mean co-morbidity scores were higher in the adherent group compared to Medicaid beneficiaries who were nonadherent to ACEI/ARBs.

Table 5: Characteristics of Medicaid patients eligible for the ACEI/ARB measure

Patient Characteristics	Prevalence(%) or Mean ± SD		
	Overall	**Adherent**	**Nonadherent**
Age	50.00 ± 9.89	51.38 ± 9.31	48.48 ± 10.27
Sex			
• Female	64.42	63.11	65.87
• Male	35.58	36.89	34.13
Race			
• White	40.53	45.36	35.21
• Black	29.02	24.82	33.63
• Hispanic	17.06	15.58	18.67
• Other	13.40	14.23	12.48
RxRisk	6.69 ± 3.00	6.95 ± 3.01	6.39 ± 2.95
RxRisk (Weighted)	8.47 ± 5.17	8.68 ± 5.22	8.25 ± 5.16
CCI	1.21 (1.23)	1.23 ± 1.20	1.18 ± 1.25
CCI (Weighted)	1.94 (1.98)	1.98 ± 1.94	1.89 ± 2.02

The results of the risk-adjusted analyses predicting adherence to ACEI/ARBs are presented in Table 6. The odds ratio estimates from the classical logistic regression model and the hierarchical logistic regression model with a random intercept are similar. All parameters included in the models, including age, sex, race/ethnicity, and RxRisk score were significant predictors of adherence with ACEI/ARBs. Models including weighted RxRisk score, CCI, and unweighted CCI instead of RxRisk were estimated and the c-statistic was comparable across the

four models. An additional model with the 42 RxRisk categories in place of the RxRisk index was analyzed, and the c-statistic was slightly higher at 0.63 compared to 0.61 for the RxRisk index and CCI models. However, 42 covariates cannot be fitted in the hierarchical regression model where the random effect only has 43 levels. Therefore, models with the RxRisk index were used for case-mix adjustment of all measures in the study. The c-statistic was 0.613 for both the classical logistic and hierarchical logistic regression models, showing modest discriminative ability for both models.

The race/ethnicity variable was found to be a strong predictor of adherence behavior in both models, with Blacks being 38% less likely to adhere to ACEI/ARB medications compared to Whites. Hispanics and other racial/ethnic groups were also less likely to adherence to ACEI/ARB medications compared to Whites. Females were 11% less likely to adhere to ACEI/ARB medications compared to males. Age and RxRisk were significant predictors, but with modest association with adherence behavior.

Table 6: Odds ratio estimates for the ACEI/ARB measure

Baseline Characteristics	Classical Logistic Regression Model*	Hierarchical Logistic Regression Random Intercept Model*
	Point Estimate	Point Estimate
Age	1.028	1.029
Sex		
• Female vs. Male	0.880	0.883
Race		
• Black vs. White	0.607	0.618
• Hispanic vs. White	0.654	0.677
• Other vs. White	0.842	0.879
RxRisk	1.044	1.040

*significant at $p < 0.0001$

In the hierarchical logistic regression model with a random-intercept, the state-level variance component was estimated to be 0.03669 (SE: 0.008275). Testing the null hypothesis of no random effects using a likelihood ratio test based on residual pseudo-likelihood yielded a chi-square of 3123.82 ($p<0.0001$) indicating the presence of a random effect. Therefore, case-mix adjusted scores based on hierarchical logistic regression model with a random intercept are presented in Table 7. The residual intra class correlation coefficient (ρ) for the random intercept model was estimated to be 0.01103 which indicates that 1.1% of the unexplained variation after controlling for patient level variables could be attributed to variation between states.

Table 7 shows the agreement between the crude and case-mix adjusted scores. The case-mix adjusted ACEI/ARB measure scores ranged from 45.5% to 65.24%, a decrease in range compared to unadjusted scores. Case-mix adjustment ranked the states differently with 12% (5) states ranked the same, 12% (5) of states changing one position in the ranking order and 77% (33) changing more than two positions. There was moderate agreement in rankings based on the crude scores and case-mix adjusted scores based on the hierarchical logistic regression model with a random intercept (Kendall's τ_b=0.80).

Table 7: Agreement in ranks: Crude and case-mix adjusted ACEI/ARB measure scores[a]

State	Unadjusted Estimates		Case-Mix Adjusted Estimates		Score difference	Rank difference
	Score (%)	Rank	Score (%)	Rank		
Alabama	52.52	**25**	53.84	**25**	-1.32	0
Alaska	-	**-**	-	**-**	-	-
Arizona	46.93	**40**	48.12	**39**	-1.19	1
Arkansas	52.72	**24**	54.99	**21**	-2.27	3
California	47.91	**38**	49.28	**37**	-1.36	1
Colorado	45.84	**41**	45.54	**43**	0.30	-2
Connecticut	61.42	**6**	60.54	**4**	0.89	2
Delaware	51.35	**30**	55.02	**19**	-3.67	11

Table 7: Agreement in ranks: Crude and case-mix adjusted ACEI/ARB measure scores[a] (continued)

State	Unadjusted Estimates		Case-Mix Adjusted Estimates		Score difference	Rank difference
	Score (%)	Rank	Score (%)	Rank		
District of Columbia	49.08	**34**	53.53	**28**	-4.45	6
Florida	55.16	**17**	54.88	**22**	0.29	-5
Georgia	45.43	**42**	46.51	**42**	-1.09	0
Hawaii	48.82	**36**	46.72	**41**	2.11	-5
Idaho	68.36	**2**	63.45	**2**	4.91	0
Illinois	52.35	**26**	55.66	**16**	-3.31	10
Indiana	50.27	**32**	48.98	**38**	1.29	-6
Iowa	60.74	**8**	58.90	**7**	1.84	1
Kansas	58.70	**12**	57.27	**11**	1.42	1
Kentucky	61.05	**7**	56.89	**12**	4.16	-5
Louisiana	49.79	**33**	52.98	**29**	-3.19	4
Maine	-	**-**	-	**-**	-	-
Maryland	49.00	**35**	52.31	**31**	-3.31	4
Massachusetts	54.24	**18**	51.70	**32**	2.55	-14
Michigan	-	**-**	-	**-**	-	-
Minnesota	51.46	**29**	51.66	**33**	-0.20	-4
Mississippi	42.10	**43**	47.01	**40**	-4.91	3
Missouri	57.97	**15**	56.78	**13**	1.20	2
Montana	-	**-**	-	**-**	-	-
Nebraska	58.87	**11**	58.32	**9**	0.55	2
Nevada	58.13	**13**	55.89	**15**	2.23	-2
New Hampshire	63.21	**4**	58.40	**8**	4.81	-4
New Jersey	52.09	**28**	53.61	**26**	-1.52	2
New Mexico	53.00	**23**	53.55	**27**	-0.55	-4
New York	53.40	**21**	54.65	**23**	-1.24	-2
North Carolina	52.20	**27**	54.52	**24**	-2.32	3
North Dakota	-	**-**	-	**-**	-	-
Ohio	-	**-**	-	**-**	-	-
Oklahoma	50.86	**31**	49.52	**36**	1.34	-5
Oregon	62.88	**5**	57.68	**10**	5.20	-5
Pennsylvania	64.53	**3**	60.76	**3**	3.77	0
Rhode Island	57.59	**16**	59.40	**6**	-1.81	10
South Carolina	48.38	**37**	52.49	**30**	-4.10	7
South Dakota	-	**-**	-	**-**	-	-

Table 7: Agreement in ranks: Crude and case-mix adjusted ACEI/ARB measure scores[a] (continued)

State	Unadjusted Estimates		Case-Mix Adjusted Estimates		Score difference	Rank difference
	Score (%)	Rank	Score (%)	Rank		
Tennessee	47.38	**39**	49.64	**35**	-2.27	4
Texas	53.73	**20**	55.43	**18**	-1.70	2
Utah	53.16	**22**	51.23	**34**	1.93	-12
Vermont	68.96	**1**	65.24	**1**	3.73	0
Virginia	58.07	**14**	59.78	**5**	-1.71	9
Washington	59.93	**10**	56.51	**14**	3.42	-4
West Virginia	60.72	**9**	55.45	**17**	5.27	-8
Wisconsin	53.77	**19**	55.00	**20**	-1.23	-1
Wyoming	-	**-**	-	**-**	-	-
[b]Kendall's τ_b	0.80					

[a]ACEI/ARB adherence scores could be generated for 43 states. Rankings ranged from 1-43.
[b]Kendall's τ_b is a nonparametric measure of association based on the number of concordances and discordances in rankings based on unadjusted and risk-adjusted scores.

Additionally, states were classified into top (20%), medium and bottom (20%) performers based on the crude and case-mix adjusted scores. Results of the agreement in the grouping based on the unadjusted and risk-adjusted groupings are shown in Table 8a and Table 8b. Results based on the two risk adjustment models showed strong agreement in grouping (κ=0.91). However, there was moderate agreement in classification of the states based on the crude and case-mix adjusted models when case-mix adjustment was conducted using both classical logistic regression (κ=0.57) and random intercept models (κ=0.65).

An alternate methodology was proposed for classification of states by identifying outliers as low (high) quality outliers if the score for a state was significantly lower (or higher) than the average score according to the 95% CI of the measure. However, the distribution of scores for the 43 states is leptokurtic i.e., concentrated about the mean, therefore not conducive to identifying outliers. The results of agreement in classification based on this methodology are included in Table A-5 (APPENDIX). Choropleth maps depicting the top, medium and bottom

performing states based on the crude and case-mix adjustment scores estimated from the hierarchical logistic regression model are depicted in Figure 1. Additional choropleth maps illustrating the distribution of crude adherence scores are presented in Figure A-3 (APPENDIX).

Table 8a: Agreement in groups: Crude and case-mix adjusted ACEI/ARB scores

Groups Based on Crude Estimates	Groups Based on Case-mix Adjusted Estimates					
	Classical Logistic Regression Model			Hierarchical Logistic Regression Random Intercept Model		
	Bottom	Medium	Top	Bottom	Medium	Top
Bottom (~20%)	6	2	0	6	2	0
Medium (60%)	2	22	3	2	23	2
Top (~20%)	0	3	5	0	2	6
Percentage misclassified[a]	25.0%	18.5%	37.5%	25.0%	14.8%	25.0%
Cohen's κ[b]	0.57			0.65		

[a]Percentage misclassified was calculated for each risk adjustment method using the classification based on the risk adjustment method as the correct classification. [b]Cohen's Kappa coefficient is a measure of agreement between unadjusted performance ratings and ratings based on the risk adjustment method with higher values indicating greater agreement

Table 8b: Agreement in groups: Case-mix adjusted ACEI/ARB scores

Groups Based on Case-mix Adjusted Estimates	Hierarchical Logistic Regression Model		
Logistic Regression Model	Bottom	Medium	Top
Bottom	8	0	0
Medium	0	26	1
Top	0	1	7
Percentage misclassified[a]	0.0%	3.7%	12.5%
Cohen's κ[b]	0.91		

[a]Percentage misclassified was calculated for each risk adjustment method using the classification based on hierarchical logistic regression model as the correct classification. [b]Cohen's Kappa coefficient is a measure of agreement between unadjusted performance ratings and ratings based on the risk adjustment method with higher values indicating greater agreement

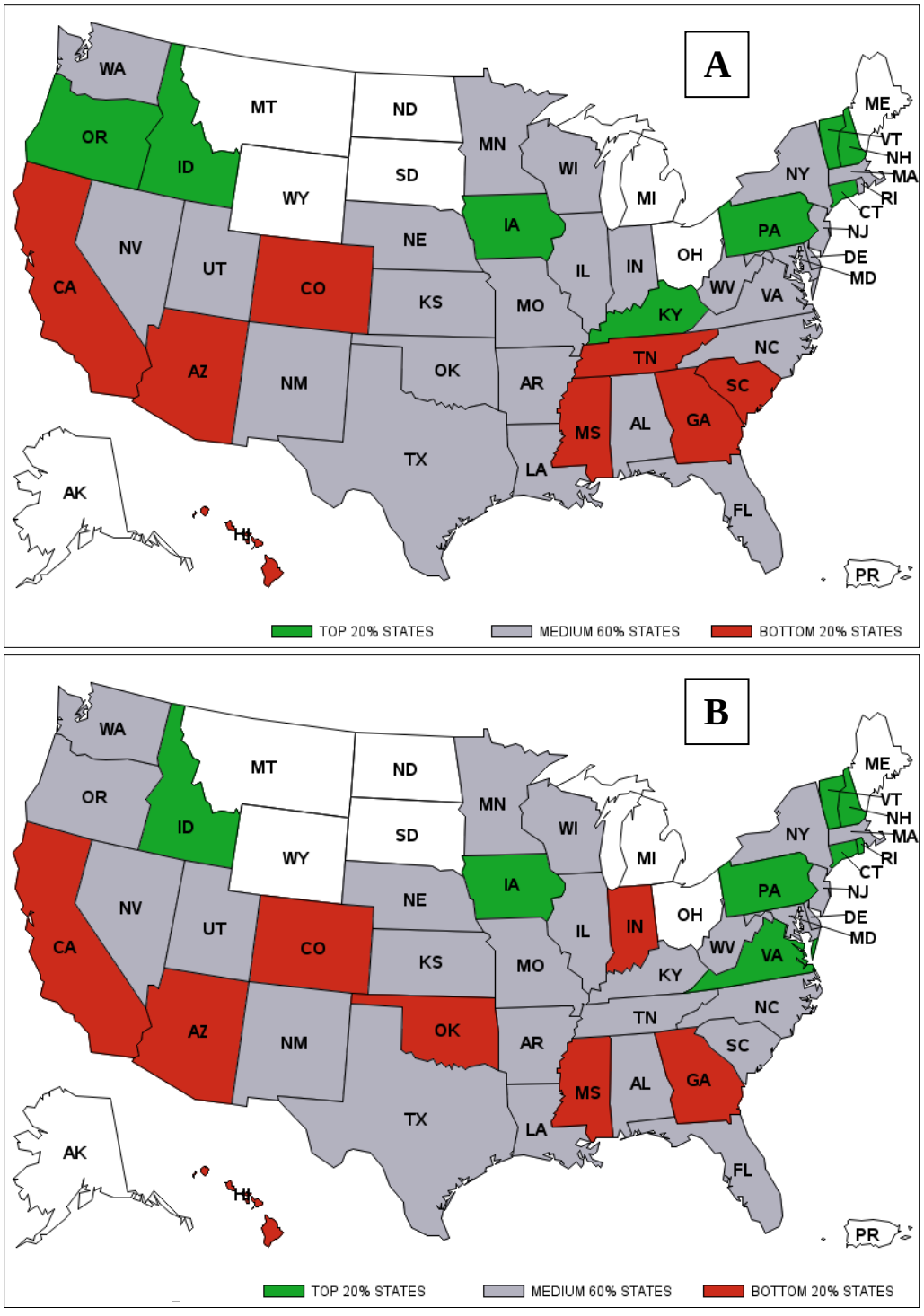

Figure 1. Interstate variations in ACEI/ARB adherence measure: (A) Unadjusted scores, (B) Case-mix adjusted scores

Beta Blockers (BB)

The PDC measure for this class of medications was computed for 42 states and the District of Columbia. The number of patients eligible for the measure in each state and the proportion of patients meeting the 80% adherence threshold are shown in Table 9. The number of patients in the denominator ranged from 969 for Idaho to 83,365 for New York. The average adherence rate for the beta blocker measure was 49.5% across all states, ranging from 34.4% of Medicaid patients adherent in Arkansas to 64.1% in Vermont. States with scores above the national benchmark are highlighted in bold.

Table 9: Patients meeting the BB measure criteria by state

State	BB	
	#patients in the denominator	%patients in the numerator
Alabama	10,715	48.81
Alaska	-	-
Arizona	12,500	44.52
Arkansas	3,774	34.42
California	63,042	46.36
Colorado	2,902	45.55
Connecticut	3,363	**59.29**
Delaware	2,030	48.23
District of Columbia	1,487	43.58
Florida	15,868	**52.19**
Georgia	12,626	43.36
Hawaii	1,514	50.00
Idaho	969	**62.33**
Illinois	20,108	**51.09**
Indiana	7,546	48.57
Iowa	3,341	**58.58**
Kansas	2,286	**55.25**
Kentucky	17,589	**56.35**
Louisiana	11,154	45.94
Maine	-	-

Table 9: Patients meeting the BB measure criteria by state (continued)

State	BB	
	#patients in the denominator	%patients in the numerator
Maryland	8,542	46.46
Massachusetts	15,274	**53.19**
Michigan	-	-
Minnesota	5,639	**50.20**
Mississippi	6,230	34.57
Missouri	10,993	**52.28**
Montana	-	-
Nebraska	1,464	**56.63**
Nevada	1,268	**56.31**
New Hampshire	1,039	**58.71**
New Jersey	9,455	**49.59**
New Mexico	3,384	**49.59**
New York	83,365	**50.47**
North Carolina	16,585	**49.64**
North Dakota	-	-
Ohio	-	-
Oklahoma	5,088	43.93
Oregon	1,210	**57.93**
Pennsylvania	7,521	**59.17**
Rhode Island	2,450	**50.45**
South Carolina	7,062	41.91
South Dakota	-	-
Tennessee	19,029	43.64
Texas	21,038	**49.73**
Utah	1,411	49.33
Vermont	2,077	**64.08**
Virginia	8,400	**54.18**
Washington	10,531	**54.48**
West Virginia	7,720	**59.65**
Wisconsin	9,263	**51.99**
Wyoming	-	-
Total	458,852	49.54

Demographic characteristics and co-morbidity measures for the eligible patient population are described in Table 10. Approximately 64% of the patients were women, and the mean age (± SD) was 49 (± 11) years. Non-adherent patients were slightly younger and a greater proportion of non-adherent patients were Black and Hispanic compared to the adherent patients. The mean co-morbidity scores were higher in the adherent group compared to Medicaid beneficiaries who were nonadherent to beta-blockers.

Table 10: Characteristics of Medicaid patients eligible for BB measure

Patient Characteristics	Prevalence(%) or Mean ± SD		
	Overall	**Adherent**	**Nonadherent**
Age	49.12 ± 10.79	50.78 ± 10.08	47.49 ± 11.20
Sex			
• Female	64.33	61.97	66.65
• Male	35.67	38.03	33.35
Race			
• White	47.30	52.00	42.68
• Black	24.96	20.52	29.31
• Hispanic	14.32	13.09	15.53
• Other	13.43	14.39	12.48
RxRisk	6.97 ± 3.10	7.23 ± 3.12	6.72 ± 3.05
RxRisk (Weighted)	8.17 ± 5.46	8.38 ± 5.51	7.97 ± 5.41
CCI	1.20 ± 1.36	1.22 ± 1.32	1.18 ± 1.39
CCI (Weighted)	1.86 ± 2.18	1.89 ± 2.12	1.82 ± 2.24

The results of the risk-adjusted analyses predicting adherence to BBs are presented in Table 11. The odds ratio estimates from the classical logistic regression model and the hierarchical logistic regression model with a random intercept are similar. All parameters included in the models, including age, sex, race/ethnicity, and RxRisk score were found to be significant predictors of adherence with beta-blockers. The c-statistic was 0.614 for both

classical logistic and hierarchical logistic regression models, showing modest discriminative ability for both models.

Race/ethnicity variable was a strong predictor of adherence behavior in both models, with Blacks being 40% less likely to adhere to beta-blocker medications compared to Whites. Hispanics and other racial/ethnic groups were also less likely to be adherent to beta-blocker medications compared to Whites. Females were 14% less likely to adhere to beta-blocker medications compared to males. Age and RxRisk were found to be significant predictors, but with modest association with adherence behavior.

Table 11: Odds ratio estimates for the BB measure

Baseline Characteristics	Classical Logistic Regression Model*	Hierarchical Logistic Regression Random Intercept Model*
	Point Estimate	**Point Estimate**
Age	1.028	1.028
Sex		
• Female vs. Male	0.855	0.860
Race		
• Black vs. White	0.582	0.598
• Hispanic vs. White	0.675	0.691
• Other vs. White	0.878	0.904
RxRisk	1.032	1.027

*significant at $p < 0.0001$

In the hierarchical logistic regression model with a random-intercept, the state-level variance component was estimated to be 0.04534 (SE: 0.01021). Testing the null hypothesis of no random effects using a likelihood ratio test based on residual pseudo-likelihood yielded a chi-square of 2324.92 ($p<0.0001$) indicating the presence of random effect. Therefore, case-mix adjusted scores based on hierarchical logistic regression model with a random intercept are presented in Table 12. The residual intra class correlation coefficient (ρ) for the random intercept

model was estimated to be 0.01359 which indicates that 1.36% of the unexplained variation after controlling for patient level variables could be attributed to variation between states.

Table 12 shows the agreement between the crude and case-mix adjusted scores. The case-mix adjusted BB measure scores ranged from 38.53% to 60.62%, a decrease in range compared to the crude estimates. Case-mix adjustment ranked the states differently with 24% (10) states ranked the same, 12% (5) of states changing one position in the ranking order and 65% (28) changing more than two positions. There was moderate agreement in rankings based on the crude scores and case-mix adjusted scores based on the hierarchical logistic regression model with a random intercept (Kendall's τ_b=0.80).

Table 12: Agreement in ranks: Crude and case-mix adjusted BB measure scores[a]

State	Unadjusted Estimates		Case-Mix Adjusted Estimates		Score difference	Rank difference
	Score (%)	Rank	Score (%)	Rank		
Alabama	48.81	**29**	50.00	**28**	-1.19	1
Alaska	-	-	-	-	-	-
Arizona	44.52	**36**	44.94	**38**	-0.42	-2
Arkansas	34.42	**43**	35.82	**43**	-1.40	0
California	46.36	**33**	46.89	**35**	-0.53	-2
Colorado	45.55	**35**	45.32	**37**	0.24	-2
Connecticut	59.29	**4**	58.49	**3**	0.81	1
Delaware	48.23	**31**	51.18	**21**	-2.95	10
District of Columbia	43.58	**39**	47.90	**32**	-4.32	7
Florida	52.19	**17**	51.85	**17**	0.34	0
Georgia	43.36	**40**	44.19	**40**	-0.83	0
Hawaii	50.00	**23**	46.95	**34**	3.05	-11
Idaho	62.33	**2**	58.60	**2**	3.74	0
Illinois	51.09	**19**	54.27	**11**	-3.18	8
Indiana	48.57	**30**	47.78	**33**	0.78	-3
Iowa	58.58	**7**	57.91	**4**	0.66	3
Kansas	55.25	**12**	54.52	**9**	0.73	3
Kentucky	56.35	**10**	52.72	**15**	3.63	-5
Louisiana	45.94	**34**	48.75	**30**	-2.82	4

Table 12: Agreement in ranks: Crude and case-mix adjusted BB measure scores[a] (cont.)

State	Unadjusted Estimates		Case-Mix Adjusted Estimates		Score difference	Rank difference
	Score (%)	Rank	Score (%)	Rank		
Maine	-	-	-	-	-	-
Maryland	46.46	**32**	49.27	**29**	-2.81	3
Massachusetts	53.19	**15**	50.66	**26**	2.53	-11
Michigan	-	-	-	-	-	-
Minnesota	50.20	**22**	50.99	**22**	-0.79	0
Mississippi	34.57	**42**	38.53	**42**	-3.96	0
Missouri	52.28	**16**	51.72	**19**	0.56	-3
Montana	-	-	-	-	-	-
Nebraska	56.63	**9**	56.81	**5**	-0.18	4
Nevada	56.31	**11**	54.00	**12**	2.31	-1
New Hampshire	58.71	**6**	55.49	**8**	3.22	-2
New Jersey	49.76	**24**	50.92	**24**	-1.16	0
New Mexico	49.59	**27**	50.22	**27**	-0.63	0
New York	50.47	**20**	50.98	**23**	-0.52	-3
North Carolina	49.64	**26**	51.79	**18**	-2.15	8
North Dakota	-	-	-	-	-	-
Ohio	-	-	-	-	-	-
Oklahoma	43.93	**37**	43.20	**41**	0.72	-4
Oregon	57.93	**8**	53.44	**14**	4.50	-6
Pennsylvania	59.17	**5**	56.20	**6**	2.96	-1
Rhode Island	50.45	**21**	51.71	**20**	-1.26	1
South Carolina	41.91	**41**	44.85	**39**	-2.93	2
South Dakota	-	-	-	-	-	-
Tennessee	43.64	**38**	45.44	**36**	-1.80	2
Texas	49.73	**25**	50.86	**25**	-1.12	0
Utah	49.33	**28**	48.08	**31**	1.25	-3
Vermont	64.08	**1**	60.62	**1**	3.46	0
Virginia	54.18	**14**	55.75	**7**	-1.57	7
Washington	54.48	**13**	52.46	**16**	2.02	-3
West Virginia	59.65	**3**	54.27	**10**	5.38	-7
Wisconsin	51.99	**18**	53.47	**13**	-1.48	5
Wyoming	-	-	-	-	-	-
[b]Kendall's τ_b	0.80					

[a]BB adherence scores could be generated for 43 states. Rankings ranged from 1-43.

[b]Kendall's τ_b is a nonparametric measure of association based on the number of concordances and discordances in rankings based on unadjusted and risk-adjusted scores

Additionally, states were classified into top (20%), medium and bottom (20%) performers based on the crude and case-mix adjusted scores. Results of the agreement in the grouping based on the unadjusted and risk-adjusted groupings are shown in Table 13a and Table 13b. Results based on the two risk adjustment models showed perfect agreement in classification of states into top, medium and bottom groups (κ=1.00). There was good agreement in classification of the states based on the crude and case-mix adjusted models when case-mix adjustment was conducted using both classical logistic regression (κ=0.74) and random intercept models (κ=0.74).

An alternate methodology was proposed for classification of states by identifying outliers as low (high) quality outliers if the score for a state was significantly lower (or higher) than the average score according to the 95% CI of the measure. However, the distribution of scores for the 43 states is leptokurtic i.e., concentrated about the mean, therefore not conducive to identifying outliers. The results of agreement in classification based on this methodology are included in Table A-6 (APPENDIX). Choropleth maps depicting the top, medium and bottom performing states based on the crude and case-mix adjustment scores estimated from the hierarchical logistic regression model are depicted in Figure 2. Additional choropleth maps illustrating the distribution of crude adherence scores are presented in Figure A-4 (APPENDIX).

Table 13a: Agreement in groups: Crude and case-mix adjusted BB measure scores

Groups Based on Crude Estimates	Groups Based on Case-mix Adjusted Estimates					
	Classical Logistic Regression Model			Hierarchical Logistic Regression Random Intercept Model		
	Bottom	Medium	Top	Bottom	Medium	Top
Bottom (~20%)	7	1	0	7	1	0
Medium (60%)	1	24	2	1	24	2
Top (~20%)	0	2	6	0	2	6
Percentage misclassified[a]	12.5%	11.1%	25%	12.5%	11.1%	25%
Cohen's κ[b]	0.74			0.74		

[a]Percentage misclassified was calculated for each risk adjustment method using the classification based on the risk adjustment method as the correct classification. [b]Cohen's Kappa coefficient is a measure of agreement between unadjusted performance ratings and ratings based on the risk adjustment method with higher values indicating greater agreement

Table 13b: Agreement in groups: Case-mix adjusted BB measure scores

Groups Based on Case-mix Adjusted Estimates	Hierarchical Logistic Regression Model		
Logistic Regression Model	Bottom	Medium	Top
Bottom	8	0	0
Medium	0	27	0
Top	0	0	8
Percentage misclassified[a]	0%	0%	0%
Cohen's κ[b]	1.00		

[a]Percentage misclassified was calculated for each risk adjustment method using the classification based on hierarchical logistic regression model as the correct classification. [b]Cohen's Kappa coefficient is a measure of agreement between unadjusted performance ratings and ratings based on the risk adjustment method with higher values indicating greater agreement

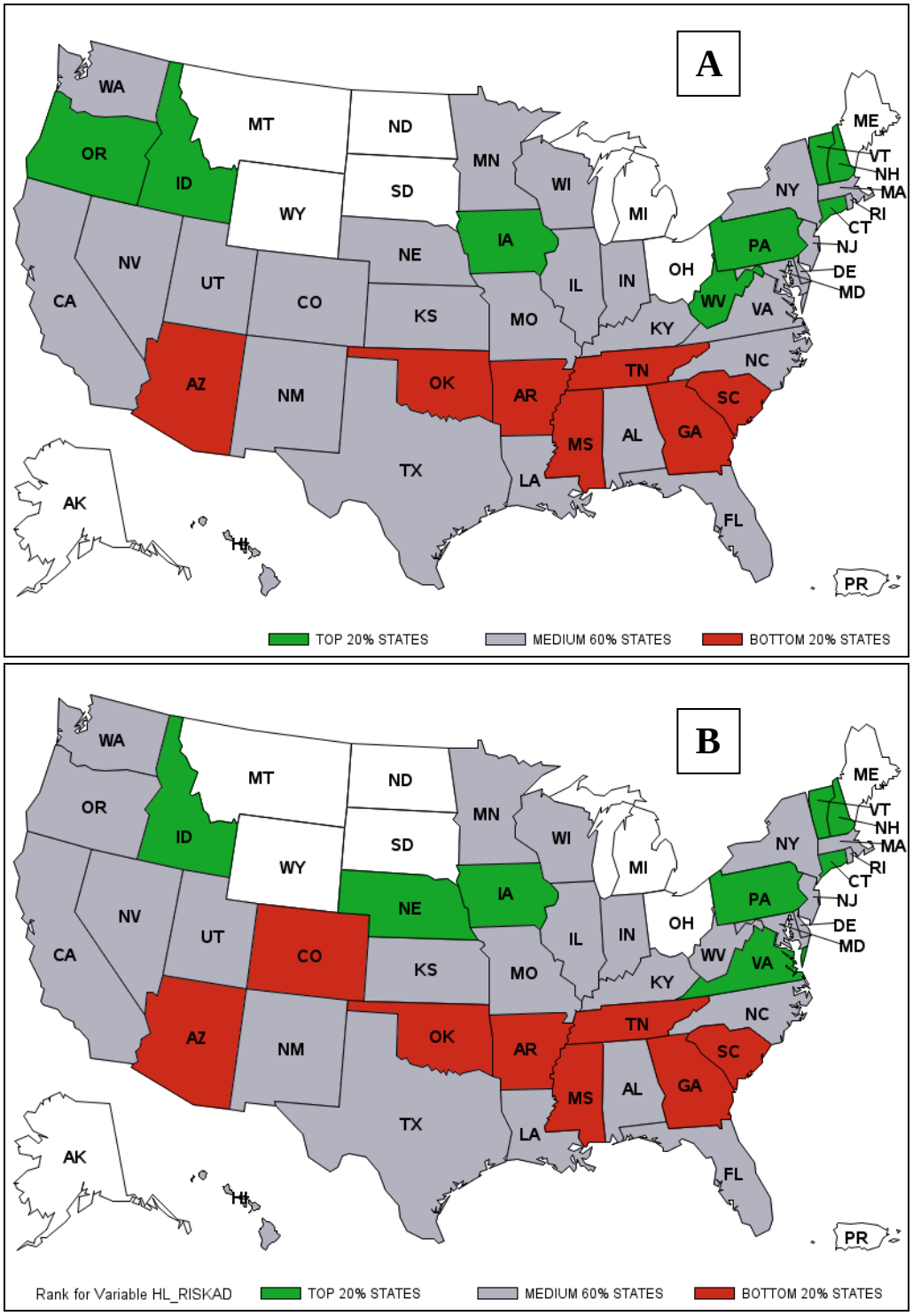

Figure 2. Interstate variations in BB adherence measure: (A) Unadjusted scores, (B) Case-mix adjusted scores

Calcium Channel Blockers (CCB)

The PDC measure for this class of medications was computed for 42 states and the District of Columbia. The number of patients eligible for the measure in each state and the proportion of patients meeting the 80% adherence threshold are shown in Table 14. The number of patients in the denominator ranged from 399 in New Hampshire to 67,257 in New York. The average adherence rate for the beta blocker measure was 49.5% across all states, ranging from 37.3% of Medicaid patients being adherent in Maryland to 69.7% in Vermont. States with scores above the national benchmark are highlighted in bold.

Table 14: Patients meeting the CCB measure criteria by state

State	CCB	
	#patients in the denominator	**%patients in the numerator**
Alabama	7,251	**52.54**
Alaska	-	-
Arizona	5,865	48.85
Arkansas	2,771	48.97
California	44,641	48.57
Colorado	1,507	45.99
Connecticut	2,207	**58.99**
Delaware	1,434	45.89
District of Columbia	1,750	49.71
Florida	12,381	**51.17**
Georgia	11,321	44.22
Hawaii	896	**52.79**
Idaho	506	**65.61**
Illinois	15,142	**51.01**
Indiana	4,640	47.91
Iowa	1,680	**58.45**
Kansas	1,503	**58.42**
Kentucky	9,563	**60.40**
Louisiana	9,712	48.93
Maine	-	-
Maryland	6,942	37.32

Table 14: Patients meeting the CCB measure criteria by state (continued)

State	CCB	
	#patients in the denominator	%patients in the numerator
Massachusetts	5,984	**55.26**
Michigan	-	-
Minnesota	2,568	48.48
Mississippi	7,027	44.51
Missouri	6,808	**52.85**
Montana	-	-
Nebraska	803	**57.04**
Nevada	1,029	**58.02**
New Hampshire	399	**64.41**
New Jersey	8,283	49.28
New Mexico	1,818	**54.79**
New York	67,257	**51.40**
North Carolina	12,962	**52.71**
North Dakota	-	-
Ohio	-	-
Oklahoma	2,527	**53.94**
Oregon	463	**60.91**
Pennsylvania	3,425	**60.79**
Rhode Island	1,616	**56.31**
South Carolina	7,159	46.04
South Dakota	-	-
Tennessee	14,004	46.87
Texas	17,109	**54.57**
Utah	873	**53.04**
Vermont	763	**69.72**
Virginia	6,179	**55.74**
Washington	4,675	**58.63**
West Virginia	3,750	**60.45**
Wisconsin	5,343	49.82
Wyoming	-	-
Total	324,536	50.94

Demographic characteristics and co-morbidity measures for the eligible patient population are described in Table 15. Approximately 67% of the patients were women, and the

mean age (± SD) was 51 ± 10 years. Non-adherent patients were slightly younger and a greater proportion of non-adherent patients were Black and Hispanic compared to adherent patients. The mean co-morbidity scores were higher in the adherent group compared to Medicaid beneficiaries' nonadherent to calcium channel blockers.

Table 15: Characteristics of Medicaid patients eligible for CCB adherence measure

Patient Characteristics	Prevalence(%) or Mean ± SD		
	Overall	**Adherent**	**Nonadherent**
Age	50.71 ± 9.87	52.29 ± 9.03	49.06 ± 10.42
Sex			
• Female	66.90	65.01	68.86
• Male	33.10	34.99	31.14
Race			
• White	33.97	38.81	28.95
• Black	38.41	33.11	43.91
• Hispanic	14.07	13.41	14.75
• Other	13.55	14.66	12.40
RxRisk	6.86 ± 3.08	7.11 ± 3.07	6.60 ± 3.06
RxRisk (Weighted)	8.12 ± 5.50	8.34 ± 5.52	7.90 ± 5.47
CCI	1.20 ± 1.31	1.22 ± 1.27	1.18 ± 1.36
CCI (Weighted)	1.89 ± 2.13	1.93 ± 2.07	1.84 ± 2.19

The results of the risk-adjusted analyses predicting adherence to CCBs are presented in Table 16. The odds ratio estimates from the classical logistic regression model and the hierarchical logistic regression model with a random intercept are similar. All parameters included in the models, including age, sex, race/ethnicity, and RxRisk score were found to be significant predictors of adherence with calcium channel blockers. The c-statistic was 0.618 for both classical logistic and hierarchical logistic regression models, showing modest discriminative ability for both models.

Race/ethnicity variable was found to be a strong predictor of adherence behavior in both models, with Blacks being 40% less likely to adhere to calcium channel blocker medications compared to Whites. Hispanics and other racial/ethnic groups were also less likely to adherence to calcium channel blocker medications compared to Whites. Females were 14% less likely to adhere to calcium channel blocker medications compared to males. Age and RxRisk were significant predictors, but with modest association with adherence behavior.

Table 16: Odds ratio estimates of patient characteristics in the risk adjustment models

Baseline Characteristics	Classical Logistic Regression Model*	Hierarchical Logistic Regression Random Intercept Model*
	Point Estimate	Point Estimate
Age	1.032	1.032
Sex		
• Female vs. Male	0.862	0.861
Race		
• Black vs. White	0.602	0.616
• Hispanic vs. White	0.674	0.689
• Other vs. White	0.833	0.871
RxRisk	1.032	1.030

*significant at $p < 0.0001$

In the hierarchical logistic regression model with a random-intercept, the state-level variance component was estimated to be 0.03238 (SE: 0.007708). Testing the null hypothesis of no random effects and complete independence of all the observations using a likelihood ratio test based on residual pseudo-likelihood yielded a chi-square of 1298.33 ($p<0.0001$) indicating the presence of random effect. Therefore, case-mix adjusted scores based on hierarchical logistic regression model with a random intercept are presented in Table 17. The residual intra class correlation coefficient (ρ) for the random intercept model was estimated to be 0.00974 which

indicates that 0.97% of the unexplained variation after controlling for patient level variables could be attributed to variation between states.

Table 17 shows the agreement between the crude and case-mix adjusted scores. The case-mix adjusted CCB measure scores ranged from 40.46% to 65.43%, a decrease in range compared to the crude estimates. Case-mix adjustment ranked the states differently with 16% (7) states ranked the same and 81% (35) changing more than two positions. There was modest agreement in rankings based on the crude scores and case-mix adjusted scores based on the hierarchical logistic regression model with a random intercept (Kendall's τ_b=0.67).

Table 17: Agreement in ranks: Crude and case-mix adjusted CCB adherence scores[a]

State	Unadjusted Estimates		Case-Mix Adjusted Estimates		Score difference	Rank difference
	Score (%)	Rank	Score (%)	Rank		
Alabama	52.54	**24**	54.66	**18**	-2.11	6
Alaska	-	**-**	-	**-**	-	-
Arizona	48.85	**33**	48.59	**39**	0.26	-6
Arkansas	48.97	**31**	51.97	**27**	-3.00	4
California	48.57	**34**	48.91	**38**	-0.34	-4
Colorado	45.99	**39**	45.72	**41**	0.26	-2
Connecticut	58.99	**8**	58.30	**4**	0.70	4
Delaware	45.89	**40**	49.99	**35**	-4.10	5
District of Columbia	49.71	**29**	53.89	**21**	-4.18	8
Florida	51.17	**26**	51.51	**28**	-0.34	-2
Georgia	44.22	**42**	45.52	**42**	-1.30	0
Hawaii	52.79	**22**	49.40	**36**	3.39	-14
Idaho	65.61	**2**	60.65	**2**	4.96	0
Illinois	51.01	**27**	54.92	**17**	-3.91	10
Indiana	47.91	**36**	48.18	**40**	-0.27	-4
Iowa	58.45	**10**	56.67	**10**	1.78	0
Kansas	58.42	**11**	56.71	**9**	1.71	2
Kentucky	60.40	**7**	55.95	**13**	4.45	-6
Louisiana	48.93	**32**	52.75	**24**	-3.83	8
Maine	-	**-**	-	**-**	-	-
Maryland	37.32	**43**	40.46	**43**	-3.13	0

Table 17: Agreement in ranks: Crude and case-mix adjusted CCB adherence scores[a] (continued)

State	Unadjusted Estimates		Case-Mix Adjusted Estimates		Score difference	Rank difference
	Score (%)	Rank	Score (%)	Rank		
Massachusetts	55.26	**16**	52.24	**26**	3.02	-10
Michigan	-	**-**	-	**-**	-	-
Minnesota	48.48	**35**	49.19	**37**	-0.71	-2
Mississippi	44.51	**41**	50.33	**34**	-5.81	7
Missouri	52.85	**21**	52.71	**25**	0.14	-4
Montana	-	**-**	-	**-**	-	-
Nebraska	57.04	**13**	57.44	**7**	-0.41	6
Nevada	58.02	**12**	56.03	**12**	1.99	0
New Hampshire	64.41	**3**	58.68	**3**	5.73	0
New Jersey	49.28	**30**	51.48	**29**	-2.20	1
New Mexico	54.79	**17**	54.10	**20**	0.69	-3
New York	51.40	**25**	52.81	**22**	-1.41	3
North Carolina	52.71	**23**	56.12	**11**	-3.42	12
North Dakota	-	**-**	-	**-**	-	-
Ohio	-	**-**	-	**-**	-	-
Oklahoma	53.94	**19**	52.76	**23**	1.17	-4
Oregon	60.91	**4**	55.58	**15**	5.33	-11
Pennsylvania	60.79	**5**	57.07	**8**	3.72	-3
Rhode Island	56.31	**14**	57.58	**6**	-1.26	8
South Carolina	46.04	**38**	50.76	**31**	-4.72	7
South Dakota	-	**-**	-	**-**	-	-
Tennessee	46.87	**37**	50.39	**33**	-3.53	4
Texas	54.57	**18**	55.88	**14**	-1.32	4
Utah	53.04	**20**	50.64	**32**	2.40	-12
Vermont	69.72	**1**	65.43	**1**	4.30	0
Virginia	55.74	**15**	58.05	**5**	-2.31	10
Washington	58.63	**9**	55.02	**16**	3.61	-7
West Virginia	60.45	**6**	54.40	**19**	6.05	-13
Wisconsin	49.82	**28**	51.44	**30**	-1.62	-2
Wyoming	-	**-**	-	**-**	-	-
[b]Kendall's τ_b	0.67					

[a]CCB adherence scores could be generated for 43 states. Rankings ranged from 1-43.

[b]Kendall's τ_b is a nonparametric measure of association based on the number of concordances and discordances in rankings based on unadjusted and risk-adjusted scores

Additionally, states were classified into top (20%), medium and bottom (20%) performers based on the crude and case-mix adjusted scores. Results of the agreement in the grouping based on the unadjusted and risk-adjusted groupings are shown in Table 18a and Table 18b. Results based on the two risk adjustment models showed perfect agreement in classification of states into top, medium and bottom groups (κ=1.00). There was poor agreement in classification of the states based on the crude and case-mix adjusted models when case-mix adjustment was conducted using both classical logistic regression (κ=0.39) and random intercept models (κ=0.39). CCB measure performed poorer compared to other PDC measures on how the crude and case-mix adjusted groupings compare.

An alternate methodology was proposed for classification of states by identifying outliers as low (high) quality outliers if the score for a state was significantly lower (or higher) than the average score according to the 95% CI of the measure. However, the distribution of scores for the 43 states is leptokurtic i.e., concentrated about the mean, therefore not conducive to identifying outliers. The results of agreement in classification based on this methodology are included in APPENDIX Table A-7. Choropleth maps depicting the top, medium and bottom performing states based on the crude and case-mix adjustment scores estimated from the hierarchical logistic regression model are depicted in Figure 3. Additional choropleth maps illustrating the distribution of crude adherence scores are presented in APPENDIX Figure A-5.

Table 18a: Agreement in groups: Crude and case-mix adjusted CCB adherence scores

Groups Based on Crude Estimates	Groups Based on Case-mix Adjusted Estimates					
	Classical Logistic Regression Model			Hierarchical Logistic Regression Random Intercept Model		
	Bottom	Medium	Top	Bottom	Medium	Top
Bottom (~20%)	4	4	0	4	4	0
Medium (60%)	4	20	3	4	20	3
Top (~20%)	0	3	5	0	3	5
Percentage misclassified[a]	50.0%	25.9%	37.5%	50.0%	25.9%	37.5%
Cohen's κ[b]	0.39			0.39		

[a]Percentage misclassified was calculated for each risk adjustment method using the classification based on the risk adjustment method as the correct classification. [b]Cohen's Kappa coefficient is a measure of agreement between unadjusted performance ratings and ratings based on the risk adjustment method with higher values indicating greater agreement

Table 18b: Agreement in groups: Case-mix adjusted CCB adherence scores

Groups Based on Case-mix Adjusted Estimates	Hierarchical Logistic Regression Model		
Logistic Regression Model	Bottom	Medium	Top
Bottom	8	0	0
Medium	0	27	0
Top	0	0	8
Percentage misclassified[a]	0%	0%	0%
Cohen's κ[b]	1.00		

[a]Percentage misclassified was calculated for each risk adjustment method using the classification based on hierarchical logistic regression model as the correct classification. [b]Cohen's Kappa coefficient is a measure of agreement between unadjusted performance ratings and ratings based on the risk adjustment method with higher values indicating greater agreement

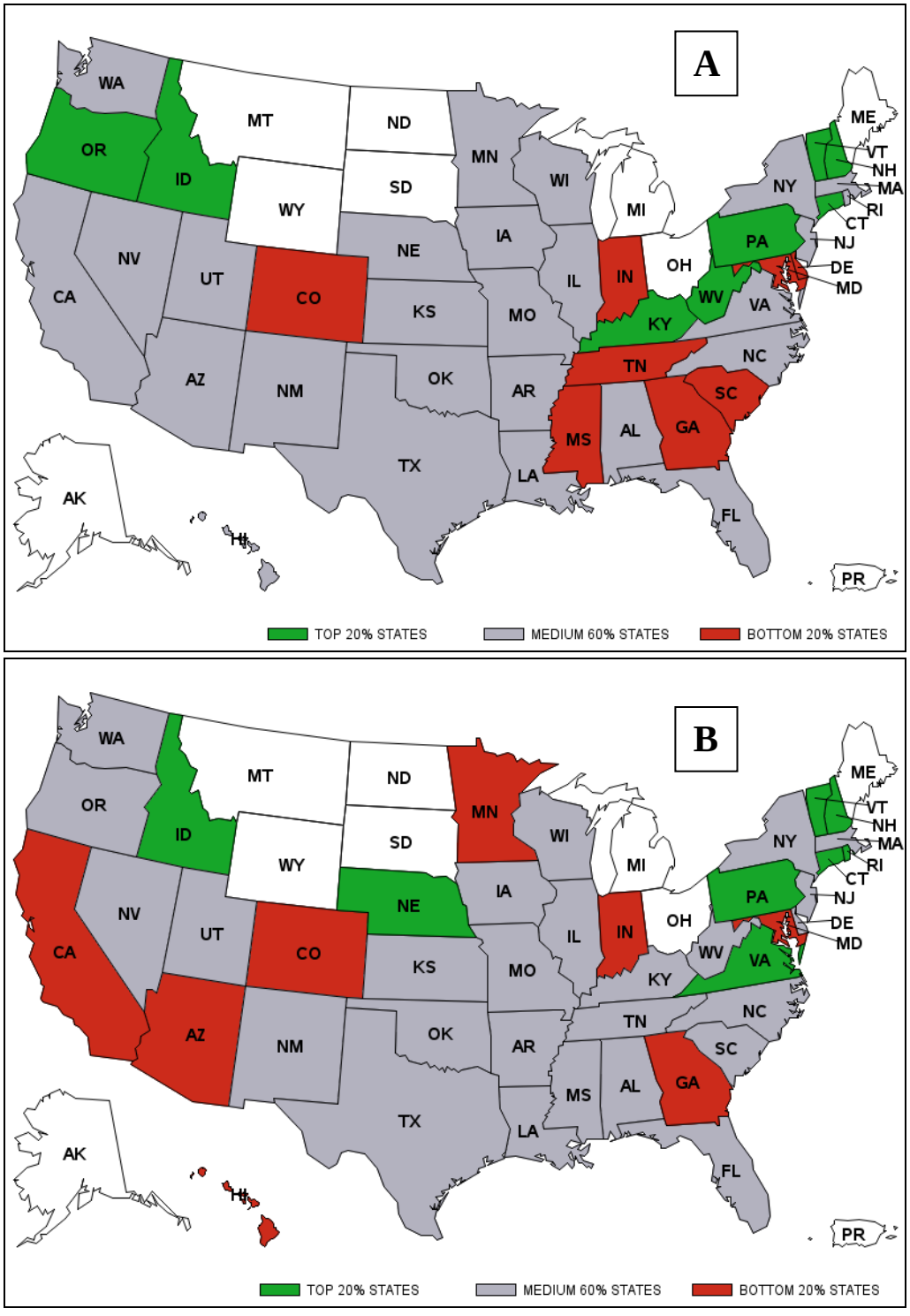

Figure 3. Interstate variations in CCB adherence measure: (A) Unadjusted scores, (B) Case-mix adjusted scores

Biguanides (BIGU)

The PDC measure for this class of medications was computed for 42 states and the District of Columbia. The number of patients eligible for the measure in each state and the proportion of patients meeting the 80% adherence threshold are tabulated in Table 19. The number of patients in the denominator ranged from 549 in New Hampshire to 65,501 in New York. The average adherence rate for the biguanide measure was 44.2% across all states, ranging from 34.1% of Medicaid patients being adherent in Mississippi to 58.3% in Vermont. Adherence rates are lower for biguanide medications compared to the antihypertensive medications included in the study. States with scores above the national benchmark are highlighted in bold.

Table 19: Patients meeting the BIGU measure criteria by state

State	BIGU	
	#patients in the denominator	**%patients in the numerator**
Alabama	6,316	44.16
Alaska	-	-
Arizona	10,246	38.59
Arkansas	2,380	40.25
California	54,942	38.03
Colorado	2,268	38.62
Connecticut	2,488	**50.36**
Delaware	1,269	42.79
District of Columbia	1,005	**44.58**
Florida	9,495	**49.83**
Georgia	7,310	39.66
Hawaii	983	**47.41**
Idaho	989	**51.67**
Illinois	12,335	**45.30**
Indiana	4,080	40.96
Iowa	1,967	**52.26**
Kansas	1,531	**47.94**
Kentucky	9,477	**52.11**
Louisiana	6,184	39.68
Maine	-	-
Maryland	5,127	43.30
Massachusetts	7,957	**46.02**

Table 19: Patients meeting the BIGU measure criteria by state (continued)

State	BIGU	
	#patients in the denominator	%patients in the numerator
Michigan	-	-
Minnesota	3,362	42.00
Mississippi	4,201	34.11
Missouri	6,497	**49.25**
Montana	-	-
Nebraska	824	**46.97**
Nevada	855	**50.06**
New Hampshire	549	**55.74**
New Jersey	7,043	**45.41**
New Mexico	3,175	44.16
New York	65,501	**47.88**
North Carolina	10,204	43.49
North Dakota	-	-
Ohio	-	-
Oklahoma	2,949	37.20
Oregon	737	**55.36**
Pennsylvania	4,069	**53.45**
Rhode Island	1,415	42.69
South Carolina	4,740	40.11
South Dakota	-	-
Tennessee	12,492	38.98
Texas	16,685	43.36
Utah	1,396	40.97
Vermont	998	**58.32**
Virginia	5,010	**50.08**
Washington	6,676	**52.50**
West Virginia	4,328	**51.83**
Wisconsin	5,615	**46.59**
Wyoming	-	-
Total	317,670	44.23

Demographic characteristics and co-morbidity measures for the eligible patient population are described in Table 20. Approximately 69% of the patients were women, and the mean age (± SD) was 49 ± 11 years. Non-adherent patients were slightly younger and a greater proportion of non-adherent patients were Black and Hispanic compared to adherent patients. The

mean co-morbidity scores were higher in the adherent group compared to Medicaid beneficiaries' nonadherent to biguanides.

Table 20: Characteristics of Medicaid patients eligible for BIGU adherence measure

Patient Characteristics	Prevalence(%) or Mean ± SD		
	Overall	**Adherent**	**Nonadherent**
Age	49.04 ± 10.50	50.92 ± 9.62	47.54 ± 10.92
Sex			
• Female	68.79	66.99	70.22
• Male	31.21	33.01	29.78
Race			
• White	38.32	43.68	34.07
• Black	23.40	19.63	26.38
• Hispanic	23.43	20.21	25.98
• Other	14.85	16.47	13.56
RxRisk	6.70 ± 3.09	7.10 ± 3.07	6.37 ± 3.06
RxRisk (Weighted)	8.71 ± 5.03	9.03 ± 5.09	8.46 ± 4.96
CCI	1.49 ± 1.02	1.52 ± 0.99	1.46 ± 1.05
CCI (Weighted)	2.58 ± 1.59	2.63 ± 1.51	2.54 ± 1.64

The results of the risk-adjusted analyses predicting adherence to biguanides are presented in Table 21. The odds ratio estimates from the classical logistic regression model and the hierarchical logistic regression model with a random intercept are similar. All parameters included in the models, including age, sex, race/ethnicity, and RxRisk score were found to be significant predictors of adherence with biguanides. The c-statistic was 0.624 for both classical logistic and hierarchical logistic regression models, showing modest discriminative ability for both models.

Race/ethnicity variable was found to be a strong predictor of adherence behavior in both models, with Blacks being 40% less likely and Hispanics 36% less likely to adhere to biguanide

medications compared to Whites. Other racial/ethnic groups were also less likely to adhere to biguanides compared to Whites. Females were 14% less likely to adhere to biguanide medications compared to males. Age and RxRisk were found to be significant predictors, but with modest association with adherence behavior.

Table 21: Odds ratio estimates of patient characteristics in the risk adjustment models

Baseline Characteristics	Classical Logistic Regression Model*	Hierarchical Logistic Regression Random Intercept Model*
	Point Estimate	Point Estimate
Age	1.029	1.029
Sex		
• Female vs. Male	0.860	0.861
Race		
• Black vs. White	0.606	0.602
• Hispanic vs. White	0.636	0.641
• Other vs. White	0.917	0.920
RxRisk	1.052	1.048

*significant at $p < 0.0001$

In the hierarchical logistic regression model with a random-intercept, the state-level variance component was estimated to be 0.02988 (SE: 0.006952). Testing the null hypothesis of no random effects and complete independence of all the observations using a likelihood ratio test based on residual pseudo-likelihood yielded a chi-square of 1723.40 ($p<0.0001$) indicating the presence of random effect. Therefore, case-mix adjusted scores based on hierarchical logistic regression model with a random intercept are presented in Table 22. The residual intra class correlation coefficient (ρ) for the random intercept model was estimated to be 0.00900 which indicates that only 0.90% of the unexplained variation after controlling for patient level variables could be attributed to variation between states.

Table 22 shows the agreement between the crude and case-mix adjusted scores. The case-mix adjusted BIGU measure scores ranged from 36% to 54%, a decrease in range compared to the crude estimates. Case-mix adjustment ranked the states differently with 16% (7) states ranked the same and 81% (35) changing more than two positions. There was good agreement in rankings based on the crude scores and case-mix adjusted scores based on the hierarchical logistic regression model with a random intercept (Kendall's τ_b=0.72).

Table 22: Agreement in ranks: Crude and case-mix adjusted BIGU adherence scores[a]

State	Unadjusted Estimates		Case-Mix Adjusted Estimates		Score difference	Rank difference
	Score (%)	Rank	Score (%)	Rank		
Alabama	44.16	**24**	44.38	**29**	-0.22	-5
Alaska	-	**-**	-	**-**	-	-
Arizona	38.59	**40**	41.59	**35**	-3.00	5
Arkansas	40.25	**34**	41.86	**33**	-1.61	1
California	38.03	**41**	41.07	**37**	-3.04	4
Colorado	38.62	**39**	39.16	**41**	-0.53	-2
Connecticut	50.36	**10**	50.40	**7**	-0.04	3
Delaware	42.79	**29**	45.45	**24**	-2.66	5
District of Columbia	44.58	**23**	49.28	**11**	-4.71	12
Florida	49.83	**13**	48.96	**12**	0.87	1
Georgia	39.66	**37**	40.12	**39**	-0.46	-2
Hawaii	47.41	**17**	44.77	**27**	2.64	-10
Idaho	51.67	**9**	51.25	**3**	0.42	6
Illinois	45.30	**22**	48.85	**13**	-3.55	9
Indiana	40.96	**33**	39.63	**40**	1.33	-7
Iowa	52.26	**6**	50.52	**6**	1.74	0
Kansas	47.94	**15**	47.17	**18**	0.78	-3
Kentucky	52.11	**7**	47.06	**19**	5.04	-12
Louisiana	39.68	**36**	41.70	**34**	-2.02	2
Maine	-	**-**	-	**-**	-	-
Maryland	43.30	**28**	46.74	**20**	-3.44	8
Massachusetts	46.02	**20**	43.24	**32**	2.78	-12
Michigan	-	**-**	-	**-**	-	-
Minnesota	42.00	**31**	43.55	**31**	-1.55	0

Table 22: Agreement in ranks: Crude and case-mix adjusted BIGU adherence scores[a] (continued)

State	Unadjusted Estimates		Case-Mix Adjusted Estimates		Score difference	Rank difference
	Score (%)	Rank	Score (%)	Rank		
Mississippi	34.11	**43**	38.00	**42**	-3.89	1
Missouri	49.25	**14**	47.39	**17**	1.86	-3
Montana	-	**-**	-	**-**	-	-
Nebraska	46.97	**18**	47.42	**16**	-0.45	2
Nevada	50.06	**12**	47.46	**15**	2.59	-3
New Hampshire	55.74	**2**	51.30	**2**	4.44	0
New Jersey	45.41	**21**	46.00	**22**	-0.59	-1
New Mexico	44.16	**25**	46.41	**21**	-2.25	4
New York	47.88	**16**	49.73	**9**	-1.86	7
North Carolina	43.49	**26**	45.23	**26**	-1.74	0
North Dakota	-	**-**	-	**-**	-	-
Ohio	-	**-**	-	**-**	-	-
Oklahoma	37.20	**42**	35.77	**43**	1.43	-1
Oregon	55.36	**3**	50.63	**5**	4.73	-2
Pennsylvania	53.45	**4**	49.70	**10**	3.75	-6
Rhode Island	42.69	**30**	44.39	**28**	-1.70	2
South Carolina	40.11	**35**	43.71	**30**	-3.60	5
South Dakota	-	**-**	-	**-**	-	-
Tennessee	38.98	**38**	41.07	**36**	-2.09	2
Texas	43.36	**27**	45.72	**23**	-2.35	4
Utah	40.97	**32**	40.85	**38**	0.12	-6
Vermont	58.32	**1**	54.37	**1**	3.95	0
Virginia	50.08	**11**	51.23	**4**	-1.15	7
Washington	52.50	**5**	49.86	**8**	2.64	-3
West Virginia	51.83	**8**	45.32	**25**	6.51	-17
Wisconsin	46.59	**19**	47.65	**14**	-1.06	5
Wyoming	-	**-**	-	**-**	-	-
[b]Kendall's τ_b	0.72					

[a]BIGU adherence scores could be generated for 43 states. Rankings ranged from 1-43.

[b]Kendall's τ_b is a nonparametric measure of association based on the number of concordances and discordances in rankings based on unadjusted and risk-adjusted scores

Additionally, states were classified into top (20%), medium and bottom (20%) performers based on the crude and case-mix adjusted scores. Results of the agreement in the

grouping based on the unadjusted and risk-adjusted groupings are shown in Table 23a and Table 23b. Results based on the two risk adjustment models showed perfect agreement in classification of states into top, medium and bottom groups (κ=1.00). There was moderate agreement in classification of the states based on the crude and case-mix adjusted models when case-mix adjustment was conducted using both classical logistic regression (κ=0.57) and random intercept models (κ=0.57).

An alternate methodology was proposed for classification of states by identifying outliers as low (high) quality outliers if the score for a state was significantly lower (or higher) than the average score according to the 95% CI of the measure. However, the distribution of scores for the 43 states is leptokurtic i.e., concentrated about the mean, therefore not conducive to identifying outliers. The results of agreement in classification based on this methodology are included in APPENDIX Table A-8. Choropleth maps depicting the top, medium and bottom performing states based on the crude and case-mix adjustment scores estimated from the hierarchical logistic regression model are depicted in Figure 4. Additional choropleth maps illustrating the distribution of crude adherence scores are presented in APPENDIX Figure A-6.

Table 23a: Agreement in groups: Crude and case-mix adjusted BIGU adherence scores

Groups Based on Crude Estimates	Groups Based on Case-mix Adjusted Estimates					
	Classical Logistic Regression Model			Hierarchical Logistic Regression Random Intercept Model		
	Bottom	Medium	Top	Bottom	Medium	Top
Bottom (~20%)	6	2	0	6	2	0
Medium (60%)	2	22	3	2	22	3
Top (~20%)	0	3	5	0	3	5
Percentage misclassified[a]	25%	18.5%	37.5%	25%	18.5%	37.5%
Cohen's κ[b]	0.57			0.57		

[a]Percentage misclassified was calculated for each risk adjustment method using the classification based on the risk adjustment method as the correct classification. [b]Cohen's Kappa coefficient is a measure of agreement between unadjusted performance ratings and ratings based on the risk adjustment method with higher values indicating greater agreement

Table 23b: Agreement in groups: Case-mix adjusted BIGU adherence scores

Groups Based on Case-mix Adjusted Estimates	Hierarchical Logistic Regression Model		
Logistic Regression Model	Bottom	Medium	Top
Bottom	8	0	0
Medium	0	27	0
Top	0	0	8
Percentage misclassified[a]	0%	0%	0%
Cohen's κ[b]	1.00		

[a]Percentage misclassified was calculated for each risk adjustment method using the classification based on hierarchical logistic regression model as the correct classification. [b]Cohen's Kappa coefficient is a measure of agreement between unadjusted performance ratings and ratings based on the risk adjustment method with higher values indicating greater agreement

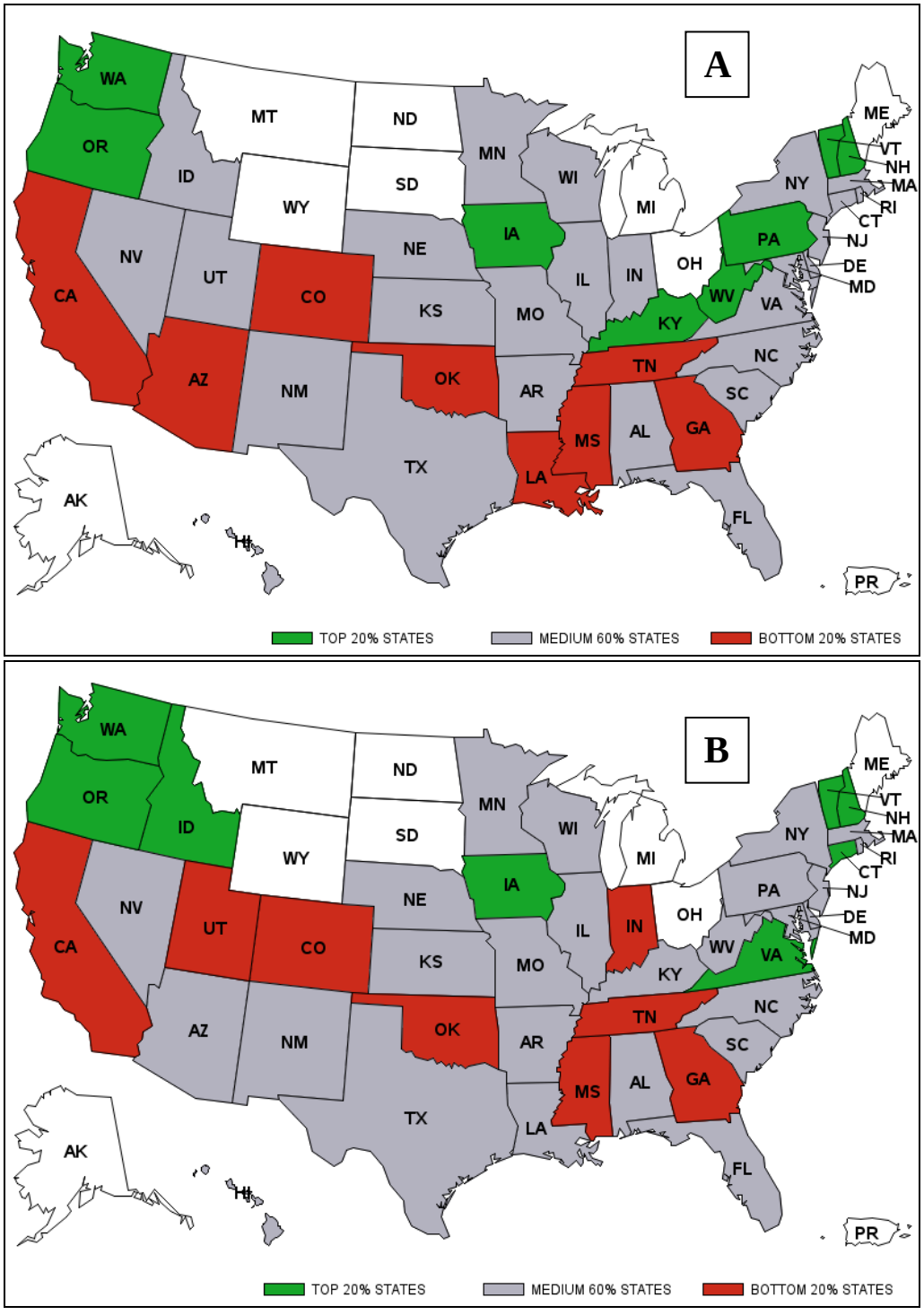

Figure 4. Interstate variations in BIGU adherence measure: (A) Unadjusted scores, (B) Case-mix adjusted scores

Sulfonylureas (SU)

The PDC measure for this class of medications was computed for 42 states and the District of Columbia. The number of patients eligible for the measure in each state and the proportion of patients meeting the 80% adherence threshold are shown in Table 24. The number of patients in the denominator ranged from 280 in New Hampshire to 40,599 in New York. The average adherence rate for the beta blocker measure was 46.6% across all states, ranging from 34.9% of Medicaid patients being adherent in Mississippi to 61.0% in Idaho. States with scores above the national benchmark are highlighted in bold.

Table 24: Patients meeting the SU measure criteria by state

State	SU	
	#patients in the denominator	%patients in the numerator
Alabama	3,957	**49.86**
Alaska	-	-
Arizona	6,165	40.16
Arkansas	1,694	43.51
California	38,842	39.49
Colorado	1,106	41.77
Connecticut	1,389	**55.08**
Delaware	720	**48.33**
District of Columbia	630	**46.67**
Florida	6,836	**50.00**
Georgia	5,031	43.35
Hawaii	739	**48.17**
Idaho	349	**61.03**
Illinois	8,477	**49.75**
Indiana	2,348	45.87
Iowa	1,088	**52.21**
Kansas	721	**47.99**
Kentucky	5,457	**56.02**
Louisiana	3,987	45.55
Maine	-	-
Maryland	3,691	46.09
Massachusetts	4,473	**48.69**
Michigan	-	-
Minnesota	1,834	44.22

Table 24: Patients meeting the SU measure criteria by state (continued)

State	SU	
	#patients in the denominator	%patients in the numerator
Mississippi	2,961	34.89
Missouri	3,467	**53.99**
Montana	-	-
Nebraska	514	**52.53**
Nevada	517	**58.41**
New Hampshire	280	**55.00**
New Jersey	4,736	**48.71**
New Mexico	2,050	45.90
New York	40,599	**49.55**
North Carolina	6,199	**48.04**
North Dakota	-	-
Ohio	-	-
Oklahoma	1,826	41.68
Oregon	390	**54.36**
Pennsylvania	2,447	**58.03**
Rhode Island	877	**47.78**
South Carolina	2,833	43.35
South Dakota	-	-
Tennessee	6,462	43.59
Texas	11,010	44.87
Utah	620	45.00
Vermont	498	**60.64**
Virginia	3,177	**54.36**
Washington	3,491	**51.93**
West Virginia	2,700	**55.33**
Wisconsin	2,994	**49.80**
Wyoming	-	-
Total	200,182	46.56

Demographic characteristics and co-morbidity measures for the eligible patient population are described in Table 25. The majority of the study population comprised of women (65%) and the mean age (± SD) was 51 ± 10 years. A greater proportion of the study population is Black or Hispanic compared to the antihypertensive medication-related PDC measures. Non-adherent patients were slightly younger and a greater proportion of non-adherent patients were

Black and Hispanic compared to adherent patients. The mean co-morbidity scores were slightly higher in the adherent group compared to Medicaid beneficiaries' nonadherent to sulfonylureas.

Table 25: Characteristics of Medicaid patients eligible for SU adherence measure

Patient Characteristics	Prevalence(%) or Mean ± SD		
	Overall	**Adherent**	**Nonadherent**
Age	50.75 ± 9.54	52.38 ± 8.71	49.32± 9.99
Sex			
• Female	65.00	63.47	66.33
• Male	35.00	36.53	33.67
Race			
• White	35.19	40.28	30.76
• Black	24.25	21.37	26.77
• Hispanic	24.31	20.55	27.58
• Other	16.24	17.80	14.89
RxRisk	6.71 ± 3.08	7.09 ± 3.06	6.38 ± 3.06
RxRisk (Weighted)	8.72 ± 5.07	9.03 ± 5.13	8.45 ± 5.00
CCI (Unweighted)	1.60 ± 1.11	1.63 ± 1.07	1.57 ± 1.13
CCI	2.77 ± 1.73	2.82 ± 1.67	2.72 ± 1.78

The results of the risk-adjusted analyses predicting adherence to sulfonylureas are presented in Table 26. The odds ratio estimates from the classical logistic regression model and the hierarchical logistic regression model with a random intercept are similar. All parameters included in the models, including age, sex, race/ethnicity, and RxRisk score were found to be significant predictors of adherence with sulfonylureas. The c-statistic was 0.620 for both classical logistic and hierarchical logistic regression models, showing modest discriminative ability for both models.

Race/ethnicity variable was found to be a strong predictor of adherence behavior in both models, with Blacks being 35% less likely to adhere to sulfonylurea medications compared to

Whites. Hispanics were 38% less likely to meet the 0.80 adherence threshold for SU medications and other racial/ethnic groups were also less likely to adhere to SU medications compared to Whites. Females were 12% less likely to adhere to SUs compared to males. Age and RxRisk were found to be significant predictors, but with modest association with adherence behavior.

Table 26: Odds ratio estimates of patient characteristics in the risk adjustment models

Baseline Characteristics	Classical Logistic Regression Model*	Hierarchical Logistic Regression Random Intercept Model*
	Point Estimate	Point Estimate
Age	1.031	1.031
Sex		
• Female vs. Male	0.878	0.890
Race		
• Black vs. White	0.650	0.646
• Hispanic vs. White	0.624	0.655
• Other vs. White	0.913	0.941
RxRisk	1.050	1.045

*significant at $p < 0.0001$

In the hierarchical logistic regression model with a random-intercept, the state-level variance component was estimated to be 0.02901 (SE: 0.007068). Testing the null hypothesis of no random effects and complete independence of all the observations using a likelihood ratio test based on residual pseudo-likelihood yielded a chi-square of 996.01 ($p<0.0001$) indicating the presence of random effect. Therefore, case-mix adjusted scores based on hierarchical logistic regression model with a random intercept are presented in Table 27. The residual intra class correlation coefficient (ρ) for the random intercept model was estimated to be 0.00874 which indicates that 0.87% of the unexplained variation after controlling for patient level variables could be attributed to variation between states.

Table 27 shows the agreement between the crude and case-mix adjusted scores. The case-mix adjusted SU measure scores ranged from 40.46% to 65.43%, a decrease in range compared to the crude estimates. Case-mix adjustment ranked the states differently with 16% (7) states ranked the same and 81% (35) changing more than two positions. There was good agreement in rankings based on the crude scores and case-mix adjusted scores based on the hierarchical logistic regression model with a random intercept (Kendall's τ_b=0.74).

Table 27: Agreement in ranks: Crude and case-mix adjusted SU adherence scores[a]

State	Unadjusted Estimates		Case-Mix Adjusted Estimates		Score difference	Rank difference
	Score (%)	Rank	Score (%)	Rank		
Alabama	49.86	**16**	50.60	**16**	-0.74	0
Alaska	-	**-**	-	**-**	-	-
Arizona	40.16	**41**	43.91	**38**	-3.75	3
Arkansas	43.51	**36**	45.47	**33**	-1.96	3
California	39.49	**42**	43.19	**40**	-3.70	2
Colorado	41.77	**39**	42.19	**41**	-0.42	-2
Connecticut	55.08	**7**	55.69	**4**	-0.61	3
Delaware	48.33	**22**	51.21	**12**	-2.87	10
District of Columbia	46.67	**27**	50.31	**19**	-3.64	8
Florida	50.00	**15**	49.93	**20**	0.07	-5
Georgia	43.35	**37**	43.83	**39**	-0.48	-2
Hawaii	48.17	**23**	45.59	**32**	2.58	-9
Idaho	61.03	**1**	56.32	**2**	4.71	-1
Illinois	49.75	**18**	53.00	**7**	-3.26	11
Indiana	45.87	**30**	44.50	**36**	1.37	-6
Iowa	52.21	**13**	50.93	**14**	1.28	-1
Kansas	47.99	**25**	47.04	**29**	0.95	-4
Kentucky	56.02	**5**	51.19	**13**	4.83	-8
Louisiana	45.55	**31**	47.12	**28**	-1.57	3
Maine	-	**-**	-	**-**	-	-
Maryland	46.09	**28**	49.74	**21**	-3.66	7
Massachusetts	48.69	**21**	46.42	**31**	2.27	-10
Michigan	-	**-**	-	**-**	-	-
Minnesota	44.22	**34**	45.40	**34**	-1.18	0

Table 27: Agreement in ranks: Crude and case-mix adjusted SU adherence scores[a] (continued)

State	Unadjusted Estimates		Case-Mix Adjusted Estimates		Score difference	Rank difference
	Score (%)	Rank	Score (%)	Rank		
Mississippi	34.89	**43**	38.68	**43**	-3.79	0
Missouri	53.99	**11**	52.75	**8**	1.24	3
Montana	-	**-**	-	**-**	-	-
Nebraska	52.53	**12**	52.15	**9**	0.38	3
Nevada	58.41	**3**	56.16	**3**	2.26	0
New Hampshire	55.00	**8**	51.31	**11**	3.69	-3
New Jersey	48.71	**20**	49.47	**24**	-0.76	-4
New Mexico	45.90	**29**	47.52	**26**	-1.62	3
New York	49.55	**19**	51.68	**10**	-2.13	9
North Carolina	48.04	**24**	49.51	**22**	-1.47	2
North Dakota	-	**-**	-	**-**	-	-
Ohio	-	**-**	-	**-**	-	-
Oklahoma	41.68	**40**	40.60	**42**	1.07	-2
Oregon	54.36	**10**	50.82	**15**	3.54	-5
Pennsylvania	58.03	**4**	54.21	**6**	3.82	-2
Rhode Island	47.78	**26**	49.18	**25**	-1.40	1
South Carolina	43.35	**38**	46.76	**30**	-3.42	8
South Dakota	-	**-**	-	**-**	-	-
Tennessee	43.59	**35**	45.31	**35**	-1.72	0
Texas	44.87	**33**	47.43	**27**	-2.56	6
Utah	45.00	**32**	44.32	**37**	0.68	-5
Vermont	60.64	**2**	57.36	**1**	3.29	1
Virginia	54.36	**9**	55.34	**5**	-0.98	4
Washington	51.93	**14**	50.34	**18**	1.59	-4
West Virginia	55.33	**6**	49.51	**23**	5.82	-17
Wisconsin	49.80	**17**	50.55	**17**	-0.75	0
Wyoming	-	**-**	-	**-**	-	-
[b]Kendall's τ_b	0.74					

[a]SU adherence scores could be generated for 43 states. Rankings ranged from 1-43.

[b]Kendall's τ_b is a nonparametric measure of association based on the number of concordances and discordances in rankings based on unadjusted and risk-adjusted scores

Additionally, states were classified into top (20%), medium and bottom (20%) performers based on the crude and case-mix adjusted scores. Results of the agreement in the

grouping based on the unadjusted and risk-adjusted groupings are shown in Table 28a and Table 28b. Results based on the two risk adjustment models showed perfect agreement in classification of states into top, medium and bottom groups (κ=1.00). There was modest agreement in classification of the states based on the crude and case-mix adjusted models when case-mix adjustment was conducted using both classical logistic regression (κ=0.57) and random intercept models (κ=0.57).

An alternate methodology was proposed for classification of states by identifying outliers as low (high) quality outliers if the score for a state was significantly lower (or higher) than the average score according to the 95% CI of the measure. However, the distribution of scores for the 43 states is leptokurtic i.e., concentrated about the mean, therefore not conducive to identifying outliers. The results of agreement in classification based on this methodology are included in APPENDIX Table A-9. Choropleth maps depicting the top, medium and bottom performing states based on the crude and case-mix adjustment scores estimated from the hierarchical logistic regression model are depicted in Figure 5. Additional choropleth maps illustrating the distribution of crude adherence scores are presented in APPENDIX Figure A-7.

Table 28a: Agreement in groups: Crude and case-mix adjusted SU adherence scores

Groups Based on Crude Estimates	Groups Based on Case-mix Adjusted Estimates					
	Classical Logistic Regression Model			Hierarchical Logistic Regression Random Intercept Model		
	Bottom	Medium	Top	Bottom	Medium	Top
Bottom (~20%)	6	2	0	6	2	0
Medium (60%)	2	22	3	2	22	3
Top (~20%)	0	3	5	0	3	5
Percentage misclassified[a]	25.0%	18.5%	37.5%	25.0%	18.5%	37.5%
Cohen's κ[b]	0.57			0.57		

[a]Percentage misclassified was calculated for each risk adjustment method using the classification based on the risk adjustment method as the correct classification. [b]Cohen's Kappa coefficient is a measure of agreement between unadjusted performance ratings and ratings based on the risk adjustment method with higher values indicating greater agreement

Table 28b: Agreement in groups: Case-mix adjusted SU adherence scores

Groups Based on Case-mix Adjusted Estimates	Hierarchical Logistic Regression Model		
Logistic Regression Model	Bottom	Medium	Top
Bottom	8	0	0
Medium	0	27	0
Top	0	0	8
Percentage misclassified[a]	0%	0%	0%
Cohen's κ[b]	1.00		

[a]Percentage misclassified was calculated for each risk adjustment method using the classification based on hierarchical logistic regression model as the correct classification. [b]Cohen's Kappa coefficient is a measure of agreement between unadjusted performance ratings and ratings based on the risk adjustment method with higher values indicating greater agreement

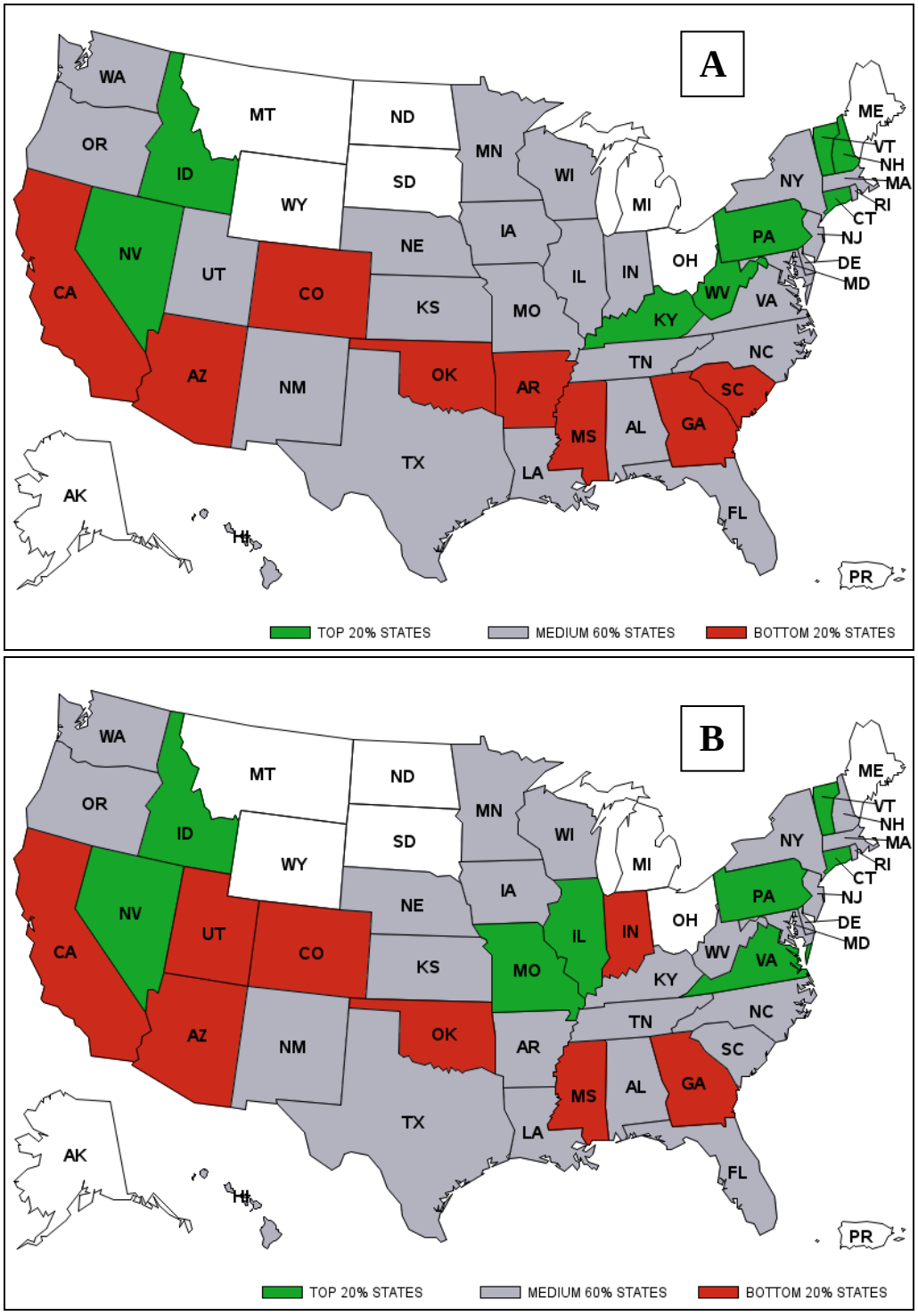

Figure 5. Interstate variations in SU adherence measure: (A) Unadjusted scores, (B) Case-mix adjusted scores

Thiazolidinediones (TZD)

The PDC measure for TZDs was computed for 42 states and the District of Columbia. The number of patients eligible for the measure in each state and the proportion of patients meeting the 80% adherence threshold are shown in Table 29. The number of patients in the denominator ranged from 242 in New Hampshire to 32,568 in New York. The average adherence rate for the beta blocker measure was 41.8% across all states, ranging from 29.7% of Medicaid patients being adherent in Mississippi to 58.8% in Vermont. States with scores above the national benchmark are highlighted in bold.

Table 29: Patients meeting the TZD measure criteria by state

State	TZD	
	#patients in the denominator	%patients in the numerator
Alabama	3,539	39.81
Alaska	-	-
Arizona	3,516	36.35
Arkansas	1,177	**44.86**
California	28,236	39.06
Colorado	946	37.63
Connecticut	1,332	**48.20**
Delaware	674	41.39
District of Columbia	360	34.72
Florida	5,584	**43.77**
Georgia	4,163	38.58
Hawaii	592	**47.47**
Idaho	509	**55.60**
Illinois	6,296	**42.44**
Indiana	1,923	40.41
Iowa	944	**52.97**
Kansas	801	**49.31**
Kentucky	4,255	**52.03**
Louisiana	3,515	41.14
Maine	-	-
Maryland	2,057	36.90
Massachusetts	3,401	**42.84**
Michigan	-	-

Table 29: Patients meeting the TZD measure criteria by state (continued)

State	TZD	
	#patients in the denominator	%patients in the numerator
Minnesota	1,356	**47.57**
Mississippi	2,118	29.70
Missouri	2,902	**49.04**
Montana	-	-
Nebraska	538	**49.63**
Nevada	415	**49.16**
New Hampshire	242	**52.48**
New Jersey	3,860	41.19
New Mexico	1,263	38.24
New York	32,568	39.65
North Carolina	5,412	**44.22**
North Dakota	-	-
Ohio	-	-
Oklahoma	1,572	**45.17**
Oregon	307	**49.51**
Pennsylvania	1,904	**55.62**
Rhode Island	712	**42.98**
South Carolina	2,493	
South Dakota	-	-
Tennessee	4,316	36.17
Texas	8,580	**43.37**
Utah	582	**43.47**
Vermont	400	**58.75**
Virginia	2,290	**47.55**
Washington	2,336	**47.39**
West Virginia	2,625	**50.93**
Wisconsin	2,358	**43.04**
Wyoming	-	-
Total	154,969	41.76

Demographic characteristics and co-morbidity measures for the eligible patient population are described in Table 30. The majority of the study population comprised of women (67%) and the mean age (± SD) was 51 ± 10 years. A greater proportion of the study population is Black or Hispanic compared to the antihypertensive medication-related PDC measures. Non-adherent patients were slightly younger, and a greater proportion of non-adherent patients were

female and of Black/Hispanic race/ethnicity compared to adherent patients. The mean RxRisk scores were slightly higher in the adherent group compared to Medicaid beneficiaries' nonadherent to thiazolidinediones. CCI weighted and unweighted indices were the same across the two groups.

Table 30: Characteristics of Medicaid patients eligible for TZD adherence measure

Patient Characteristics	Prevalence(%) or Mean ± SD		
	Overall	**Adherent**	**Nonadherent**
Age	50.93 ± 9.28	52.14 ± 8.70	50.06 ± 9.59
Sex			
• Female	67.14	64.51	69.02
• Male	32.86	35.49	30.98
Race			
• White	37.53	43.25	33.43
• Black	22.86	19.48	25.28
• Hispanic	23.72	20.12	26.29
• Other	15.89	17.15	14.99
RxRisk	7.06 ± 3.09	7.32 ± 3.06	6.89 ± 3.09
RxRisk (Weighted)	9.00 ± 5.14	9.22 ± 5.18	8.84 ± 5.11
CCI (Unweighted)	1.59 ± 1.05	1.59 ± 1.01	1.60 ± 1.08
CCI	2.75 ± 1.60	2.76 ± 1.53	2.75 ± 1.64

The results of the risk-adjusted analyses predicting adherence to thiazolidinediones are presented in Table 31. The odds ratio estimates from the classical logistic regression model and the hierarchical logistic regression model with a random intercept are similar. All parameters included in the models, including age, sex, race/ethnicity, and RxRisk score were found to be significant predictors of the TZD adherence measure. The c-statistic was 0.599 for both classical logistic and hierarchical logistic regression models, showing weak discriminative ability for both models, compared to the other measures included in the study.

Race/ethnicity and gender were strong predictors of adherence behavior in both models. Blacks were 37% less likely to adhere to TZDs compared to Whites in the classical logistic regression model. Hispanics were 39% less likely to meet the 0.80 adherence threshold for TZDs and other racial/ethnic groups were also less likely to be adherent compared to Whites. Females were 19% less likely to adhere to TZDs compared to males. Age and RxRisk were found to be significant predictors, but with modest association with adherence behavior. Similar results were observed using the hierarchical logistic regression model with state included as a random intercept.

Table 31: Odds ratio estimates of patient characteristics in the risk adjustment models

Baseline Characteristics	Classical Logistic Regression Model*	Hierarchical Logistic Regression Random Intercept Model*
	Point Estimate	Point Estimate
Age	1.023	1.023
Sex		
• Female vs. Male	0.808	0.890
Race		
• Black vs. White	0.629	0.655
• Hispanic vs. White	0.614	0.643
• Other vs. White	0.869	0.910
RxRisk	1.031	1.027

*significant at $p < 0.0001$

In the hierarchical logistic regression model with a random-intercept, the state-level variance component was estimated to be 0.03249 (SE: 0.008010). Testing the null hypothesis of no random effects and complete independence of all the observations using a likelihood ratio test based on residual pseudo-likelihood yielded a chi-square of 488.74 ($p<0.0001$) indicating the presence of random effect. Therefore, case-mix adjusted scores based on hierarchical logistic regression model with a random intercept are presented in Table 32. The residual intra class

correlation coefficient (ρ) for the random intercept model was estimated to be 0.00978 which indicates that 0.98% of the unexplained variation after controlling for patient level variables could be attributed to variation between states.

Table 32 shows the agreement between the crude and case-mix adjusted scores. The case-mix adjusted SU measure scores ranged from 40.46% to 65.43%, a decrease in range compared to the crude estimates. Case-mix adjustment ranked the states differently with 16% (7) states ranked the same and 81% (35) changing more than two positions. There was good agreement in rankings based on the crude scores and case-mix adjusted scores based on the hierarchical logistic regression model with a random intercept (Kendall's τ_b=0.78).

Table 32: Agreement in ranks: Crude and case-mix adjusted TZD adherence scores[a]

State	Unadjusted Estimates		Case-Mix Adjusted Estimates		Score difference	Rank difference
	Score (%)	Rank	Score (%)	Rank		
Alabama	39.81	**33**	40.98	**33**	-1.17	0
Alaska	-	**-**	-	**-**	-	-
Arizona	36.35	**40**	39.11	**37**	-2.76	3
Arkansas	44.86	**19**	46.71	**13**	-1.85	6
California	39.06	**35**	41.70	**31**	-2.64	4
Colorado	37.63	**38**	38.18	**40**	-0.55	-2
Connecticut	48.20	**13**	49.90	**6**	-1.70	7
Delaware	41.39	**28**	44.45	**23**	-3.06	5
District of Columbia	34.72	**42**	37.85	**42**	-3.13	0
Florida	43.77	**21**	44.29	**24**	-0.52	-3
Georgia	38.58	**36**	39.46	**35**	-0.88	1
Hawaii	47.47	**16**	45.62	**18**	1.85	-2
Idaho	55.60	**3**	53.77	**2**	1.83	1
Illinois	42.44	**27**	44.97	**20**	-2.54	7
Indiana	40.41	**32**	39.04	**39**	1.37	-7
Iowa	52.97	**4**	51.70	**3**	1.27	1
Kansas	49.31	**10**	48.51	**8**	0.80	2
Kentucky	52.03	**6**	47.86	**12**	4.17	-6
Louisiana	41.14	**30**	43.19	**28**	-2.06	2

Table 32: Agreement in ranks: Crude and case-mix adjusted TZD adherence scores[a] (cont.)

State	Unadjusted Estimates		Case-Mix Adjusted Estimates		Score difference	Rank difference
	Score (%)	Rank	Score (%)	Rank		
Maine	-	-	-	-	-	-
Maryland	36.90	**39**	39.11	**38**	-2.21	1
Massachusetts	42.84	**26**	40.88	**34**	1.96	-8
Michigan	-	-	-	-	-	-
Minnesota	47.57	**14**	47.98	**10**	-0.41	4
Mississippi	29.70	**43**	33.22	**43**	-3.52	0
Missouri	49.04	**12**	46.68	**14**	2.36	-2
Montana	-	-	-	-	-	-
Nebraska	49.63	**8**	49.99	**5**	-0.36	3
Nevada	49.16	**11**	47.96	**11**	1.20	0
New Hampshire	52.48	**5**	48.43	**9**	4.05	-4
New Jersey	41.19	**29**	42.05	**30**	-0.86	-1
New Mexico	38.24	**37**	39.32	**36**	-1.08	1
New York	39.65	**34**	41.35	**32**	-1.70	2
North Carolina	44.22	**20**	46.26	**16**	-2.04	4
North Dakota	-	-	-	-	-	-
Ohio	-	-	-	-	-	-
Oklahoma	45.17	**18**	44.18	**26**	0.99	-8
Oregon	49.51	**9**	44.92	**21**	4.59	-12
Pennsylvania	55.62	**2**	51.65	**4**	3.97	-2
Rhode Island	42.98	**25**	45.11	**19**	-2.13	6
South Carolina	40.95	**31**	44.19	**25**	-3.24	6
South Dakota	-	-	-	-	-	-
Tennessee	36.17	**41**	37.94	**41**	-1.77	0
Texas	43.37	**23**	46.65	**15**	-3.28	8
Utah	43.47	**22**	42.41	**29**	1.06	-7
Vermont	58.75	**1**	54.61	**1**	4.14	0
Virginia	47.55	**15**	48.51	**7**	-0.96	8
Washington	47.39	**17**	44.52	**22**	2.87	-5
West Virginia	50.93	**7**	46.18	**17**	4.75	-10
Wisconsin	43.04	**24**	43.63	**27**	-0.58	-3
Wyoming	-	-	-	-	-	-
[b]Kendall's τ_b	0.78					

[a]TZD adherence scores could be generated for 43 states. Rankings ranged from 1-43

[b]Kendall's τ_b is a nonparametric measure of association based on the number of concordances and discordances in rankings based on unadjusted and risk-adjusted scores

Additionally, states were classified into top (20%), medium and bottom (20%) performers based on the crude and case-mix adjusted scores. Results of the agreement in the grouping based on the unadjusted and risk-adjusted groupings are shown in Table 33a and Table 33b. Results based on the two risk adjustment models showed close to perfect agreement in classification of states into top, medium and bottom groups (κ=0.91). There was modest agreement in classification of the states based on the crude and case-mix adjusted models when case-mix adjustment was conducted using both classical logistic regression (κ=0.65) and random intercept models (κ=0.65).

An alternate methodology was proposed for classification of states by identifying outliers as low (high) quality outliers if the score for a state was significantly lower (or higher) than the average score according to the 95% CI of the measure. However, the distribution of scores for the 43 states is leptokurtic i.e., concentrated about the mean, therefore not conducive to identifying outliers. The results of agreement in classification based on this methodology are included in APPENDIX Table A-10. Choropleth maps depicting the top, medium and bottom performing states based on the crude and case-mix adjustment scores estimated from the hierarchical logistic regression model are depicted in Figure 6. Additional choropleth maps illustrating the distribution of crude adherence scores are presented in APPENDIX Figure A-8.

Table 33a: Agreement in groups: Crude and case-mix adjusted TZD adherence scores

Groups Based on Crude Estimates	Groups Based on Case-mix Adjusted Estimates					
	Classical Logistic Regression Model			Hierarchical Logistic Regression Random Intercept Model		
	Bottom	Medium	Top	Bottom	Medium	Top
Bottom (~20%)	7	1	0	7	1	0
Medium (60%)	1	23	3	1	23	3
Top (~20%)	0	3	5	0	3	5
Percentage misclassified[a]	12.5%	14.8%	37.5%	12.5%	14.8%	37.5%
Cohen's κ[b]	0.65			0.65		

[a]Percentage misclassified was calculated for each risk adjustment method using the classification based on the risk adjustment method as the correct classification. [b]Cohen's Kappa coefficient is a measure of agreement between unadjusted performance ratings and ratings based on the risk adjustment method with higher values indicating greater agreement

Table 33b: Agreement in groups: Case-mix adjusted TZD adherence scores

Groups Based on Case-mix Adjusted Estimates	Hierarchical Logistic Regression Model		
Logistic Regression Model	Bottom	Medium	Top
Bottom	7	1	0
Medium	1	26	0
Top	0	0	8
Percentage misclassified[a]	12.5%	3.7%	0%
Cohen's κ[b]	0.91		

[a]Percentage misclassified was calculated for each risk adjustment method using the classification based on hierarchical logistic regression model as the correct classification. [b]Cohen's Kappa coefficient is a measure of agreement between unadjusted performance ratings and ratings based on the risk adjustment method with higher values indicating greater agreement.

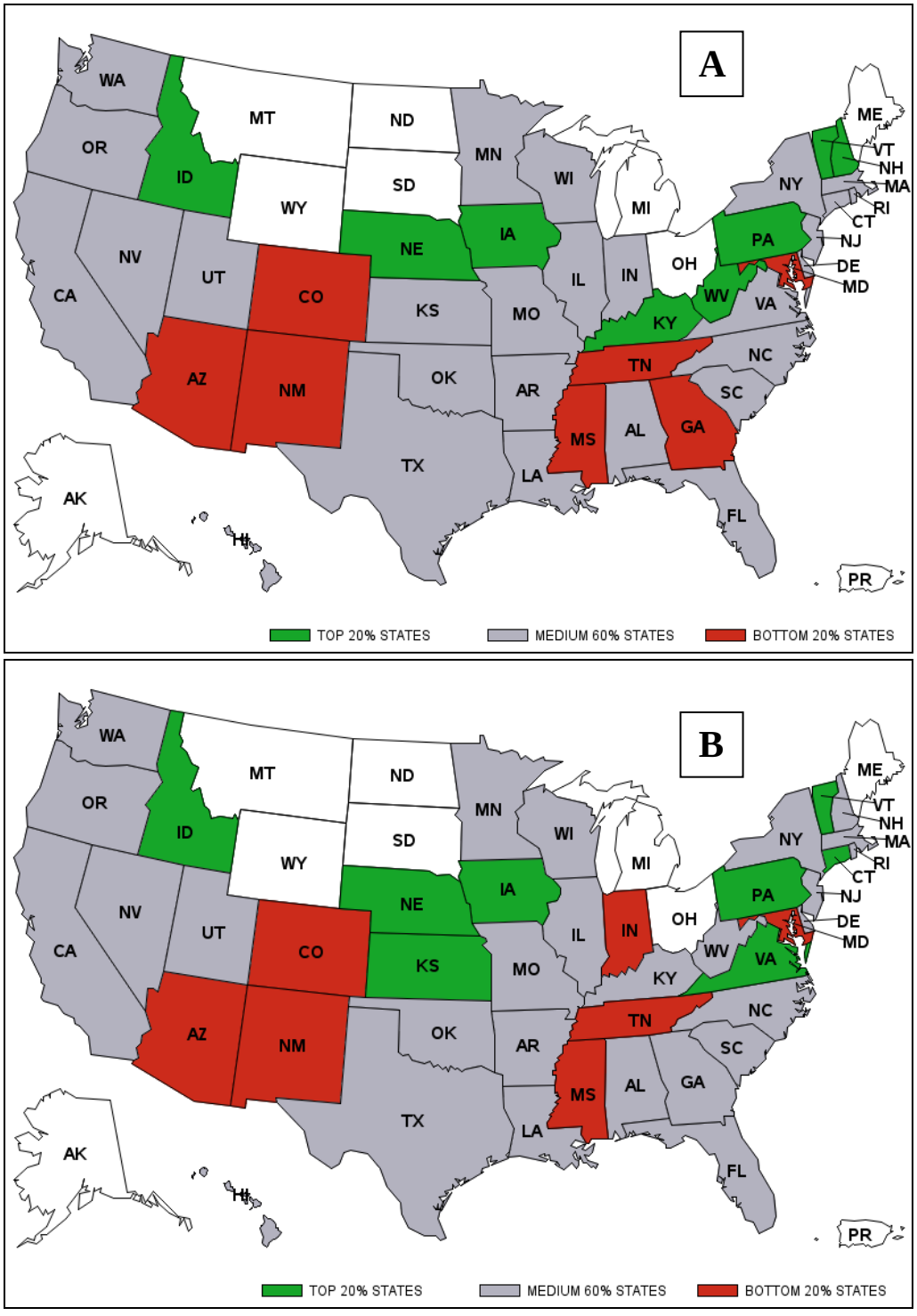

Figure 6. Interstate variations in TZD adherence measure: (A) Unadjusted scores, (B) Case-mix adjusted scores

Statins (STAT)

The PDC measure for this class of medications was computed for 42 states and the District of Columbia. The number of patients eligible for the measure in each state and the proportion of patients meeting the 80% adherence threshold are shown in Table 34. The number of patients in the denominator ranged from 1,182 in Oregon to 121,522 in New York. The average adherence rate for the statin adherence measure was 48.3% across all states, ranging from 38.4% of Medicaid patients being adherent in Mississippi to 67.6% in Vermont. States with scores above the national benchmark are highlighted in bold.

Table 34: Patients meeting the STAT measure criteria by state

State	STAT	
	#patients in the denominator	**%patients in the numerator**
Alabama	11,258	**51.02**
Alaska	-	-
Arizona	14,405	41.66
Arkansas	3,776	46.00
California	84,254	42.78
Colorado	3,437	44.57
Connecticut	4,652	**58.19**
Delaware	2,664	47.18
District of Columbia	1,843	45.25
Florida	19,897	**51.19**
Georgia	13,710	44.35
Hawaii	2,260	**54.47**
Idaho	1,201	**66.19**
Illinois	21,752	**49.17**
Indiana	7,702	**49.51**
Iowa	3,495	**61.34**
Kansas	2,653	**58.20**
Kentucky	20,199	**56.84**
Louisiana	11,396	47.26
Maine	-	-
Maryland	9,277	47.03
Massachusetts	16,860	46.79
Michigan	-	-
Minnesota	5,634	**50.75**

Table 34: Patients meeting the STAT measure criteria by state (continued)

State	STAT	
	#patients in the denominator	%patients in the numerator
Mississippi	6,165	38.44
Missouri	11,733	**57.74**
Montana	-	-
Nebraska	1,496	**59.83**
Nevada	1,581	**54.59**
New Hampshire	1,224	**61.27**
New Jersey	12,877	45.25
New Mexico	3,889	**48.78**
New York	121,522	46.21
North Carolina	18,879	**51.70**
North Dakota	-	-
Ohio	-	-
Oklahoma	4,854	**52.51**
Oregon	1,182	**59.31**
Pennsylvania	8,559	**58.11**
Rhode Island	3,242	43.58
South Carolina	8,646	45.95
South Dakota	-	-
Tennessee	19,197	41.24
Texas	25,742	**50.91**
Utah	1,882	**51.97**
Vermont	2,303	**67.56**
Virginia	9,334	**57.67**
Washington	11,460	**57.23**
West Virginia	9,968	**56.99**
Wisconsin	9,705	**52.50**
Wyoming	-	-
Total	557,765	48.32

Demographic characteristics and co-morbidity measures for the eligible patient population are described in Table 35. The majority of the study population was comprised of women (64%) and the mean age (± SD) was 51 ± 9 years. A greater proportion of the study population is Black or Hispanic compared to the antihypertensive medication-related PDC measures. Non-adherent patients were slightly younger and a greater proportion of non-adherent patients were women, of Black and Hispanic race/ethnicity compared to adherent patients. The

mean comorbidity scores were slightly higher in the adherent group compared to Medicaid beneficiaries' nonadherent to statins.

Table 35: Characteristics of Medicaid patients eligible for STAT adherence measure

Patient Characteristics	Prevalence(%) or Mean ± SD		
	Overall	**Adherent**	**Nonadherent**
Age	51.37 ± 8.98	52.39 ± 8.63	50.41 ± 9.20
Sex			
• Female	64.12	62.48	65.66
• Male	35.88	37.52	34.34
Race			
• White	45.79	51.61	40.35
• Black	20.07	17.32	22.64
• Hispanic	18.08	14.49	21.42
• Other	16.06	16.57	15.59
RxRisk	6.95 ± 3.07	7.27 ± 3.07	6.65 ± 3.04
RxRisk (Weighted)	7.65 ± 5.41	8.09 ± 5.45	7.23 ± 5.34
CCI (Unweighted)	1.18 ± 1.21	1.21 ± 1.20	1.14 ± 1.22
CCI	1.87 ± 1.90	1.93 ± 1.88	1.81 ± 1.92

The results of the risk-adjusted analyses predicting adherence to statins are presented in Table 36. The odds ratio estimates from the classical logistic regression model and the hierarchical logistic regression model with a random intercept are similar. All parameters included in the models, including age, sex, race/ethnicity, and RxRisk score were found to be significant predictors of the STAT adherence measure. The c-statistic was 0.611 for both classical logistic and hierarchical logistic regression models, showing weak discriminative ability for both models, compared to the other measures included in the study.

Race/ethnicity and gender were found to be strong predictors of adherence behavior in both models. Blacks were 40% less likely, Hispanics were 45% less likely, and other

racial/ethnic groups were 17% less likely to meet the 0.80 adherence threshold for statins. Females were 18% less likely to adhere to statins compared to males. Age and RxRisk were found to be significant predictors, but with modest association with adherence behavior.

Table 36: Odds ratio estimates of patient characteristics in the risk adjustment models

Baseline Characteristics	Classical Logistic Regression Model*	Hierarchical Logistic Regression Random Intercept Model*
	Point Estimate	Point Estimate
Age	1.025	1.025
Sex		
• Female vs. Male	0.819	0.820
Race		
• Black vs. White	0.609	0.619
• Hispanic vs. White	0.550	0.569
• Other vs. White	0.832	0.864
RxRisk	1.057	1.051

*significant at $p < 0.0001$

In the hierarchical logistic regression model with a random-intercept, the state-level variance component was estimated to be 0.04566 (SE: 0.01027). Testing the null hypothesis of no random effects using a likelihood ratio test based on residual pseudo-likelihood yielded a chi-square of 3095.49 ($p<0.0001$) indicating the presence of random effect. Therefore, case-mix adjusted scores based on hierarchical logistic regression model with a random intercept are presented in Table 37. The residual intra class correlation coefficient (ρ) for the random intercept model was estimated to be 0.01369 which indicates that 1.37% of the unexplained variation after controlling for patient level variables could be attributed to variation between states.

Table 37 shows the agreement between the crude and case-mix adjusted scores. The case-mix adjusted statin measure scores ranged from 42.10% to 64.05%, a decrease in range compared to the crude estimates. Case-mix adjustment ranked the states differently with 12% (5)

states ranked the same, 30% (13) states changing one position and 58% (25) changing more than two positions. There was good agreement in rankings based on the crude scores and case-mix adjusted scores based on the hierarchical logistic regression model with a random intercept (Kendall's τ_b=0.81).

Table 37: Agreement in ranks: Crude and case-mix adjusted STAT adherence scores[a]

State	Unadjusted Estimates		Case-Mix Adjusted Estimates		Score difference	Rank difference
	Score (%)	Rank	Score (%)	Rank		
Alabama	51.02	**22**	50.98	**23**	0.04	-1
Alaska	-	**-**	-	**-**	-	-
Arizona	41.66	**41**	44.44	**40**	-2.79	1
Arkansas	46.00	**33**	46.80	**34**	-0.80	-1
California	42.78	**40**	45.84	**37**	-3.06	3
Colorado	44.57	**37**	45.24	**39**	-0.66	-2
Connecticut	58.19	**8**	60.47	**3**	-2.28	5
Delaware	47.18	**29**	49.55	**30**	-2.37	-1
District of Columbia	45.25	**35**	50.16	**27**	-4.91	8
Florida	51.19	**21**	51.72	**21**	-0.52	0
Georgia	44.35	**38**	44.38	**41**	-0.03	-3
Hawaii	54.47	**16**	53.39	**15**	1.08	1
Idaho	66.19	**2**	62.09	**2**	4.11	0
Illinois	49.17	**26**	51.91	**19**	-2.74	7
Indiana	49.51	**25**	47.12	**33**	2.39	-8
Iowa	61.34	**3**	58.92	**4**	2.42	-1
Kansas	58.20	**7**	56.30	**8**	1.90	-1
Kentucky	56.84	**14**	52.49	**17**	4.36	-3
Louisiana	47.26	**28**	48.43	**32**	-1.17	-4
Maine	-	**-**	-	**-**	-	-
Maryland	47.03	**30**	49.66	**28**	-2.63	2
Massachusetts	46.79	**31**	45.36	**38**	1.43	-7
Michigan	-	**-**	-	**-**	-	-
Minnesota	50.75	**24**	50.84	**24**	-0.10	0
Mississippi	38.44	**43**	42.18	**42**	-3.74	1
Missouri	57.74	**10**	55.45	**9**	2.30	1

Table 37: Agreement in ranks: Crude and case-mix adjusted STAT adherence scores[a] (continued)

State	Unadjusted Estimates		Case-Mix Adjusted Estimates		Score difference	Rank difference
	Score (%)	Rank	Score (%)	Rank		
Montana	-	-	-	-	-	-
Nebraska	59.83	**5**	58.88	**5**	0.95	0
Nevada	54.59	**15**	52.92	**16**	1.66	-1
New Hampshire	61.27	**4**	57.18	**6**	4.09	-2
New Jersey	45.25	**36**	46.79	**35**	-1.54	1
New Mexico	48.78	**27**	50.63	**25**	-1.85	2
New York	46.21	**32**	49.56	**29**	-3.36	3
North Carolina	51.70	**20**	52.22	**18**	-0.52	2
North Dakota	-	-	-	-	-	-
Ohio	-	-	-	-	-	-
Oklahoma	52.51	**17**	50.56	**26**	1.96	-9
Oregon	59.31	**6**	54.69	**10**	4.61	-4
Pennsylvania	58.11	**9**	54.66	**11**	3.45	-2
Rhode Island	43.58	**39**	46.57	**36**	-2.99	3
South Carolina	45.95	**34**	48.44	**31**	-2.48	3
South Dakota	-	-	-	-	-	-
Tennessee	41.24	**42**	42.10	**43**	-0.86	-1
Texas	50.91	**23**	54.25	**12**	-3.34	11
Utah	51.97	**19**	50.99	**22**	0.97	-3
Vermont	67.56	**1**	64.05	**1**	3.52	0
Virginia	57.67	**11**	57.14	**7**	0.53	4
Washington	57.23	**12**	54.06	**13**	3.18	-1
West Virginia	56.99	**13**	51.80	**20**	5.19	-7
Wisconsin	52.50	**18**	53.46	**14**	-0.96	4
Wyoming	-	-	-	-	-	-
[b]Kendall's τ_b	0.81					

[a]STAT adherence scores could be generated for 43 states. Rankings ranged from 1-43

[b]Kendall's τ_b is a nonparametric measure of association based on the number of concordances and discordances in rankings based on unadjusted and risk-adjusted scores

Additionally, states were classified into top (20%), medium and bottom (20%) performers based on the crude and case-mix adjusted scores. Results of the agreement in the grouping based on the unadjusted and risk-adjusted groupings are shown in Table 38a and Table

38b. Results based on the two risk adjustment models showed close to perfect agreement in classification of states into top, medium and bottom groups (κ=0.91). There was good agreement in classification of the states based on the crude and case-mix adjusted models when case-mix adjustment was conducted using both classical logistic regression (κ=0.74) and random intercept models (κ=0.83).

An alternate methodology was proposed for classification of states by identifying outliers as low (high) quality outliers if the score for a state was significantly lower (or higher) than the average score according to the 95% CI of the measure. However, the distribution of scores for the 43 states is leptokurtic i.e., concentrated about the mean, therefore not conducive to identifying outliers. The results of agreement in classification based on this methodology are included in APPENDIX Table A-11. Choropleth maps depicting the top, medium and bottom performing states based on the crude and case-mix adjustment scores estimated from the hierarchical logistic regression model are depicted in Figure 7. Additional choropleth maps illustrating the distribution of crude adherence scores are presented in APPENDIX Figure A-9.

Table 38a: Agreement in groups: Crude and case-mix adjusted STAT adherence scores

Groups Based on Crude Estimates	Groups Based on Case-mix Adjusted Estimates					
	Classical Logistic Regression Model			Hierarchical Logistic Regression Random Intercept Model		
	Bottom	Medium	Top	Bottom	Medium	Top
Bottom (~20%)	6	2	0	7	1	0
Medium (60%)	2	24	1	1	25	1
Top (~20%)	0	1	7	0	1	7
Percentage misclassified[a]	25.0%	11.1%	12.5%	12.5%	7.4%	12.5%
Cohen's κ[b]	0.74			0.83		

[a]Percentage misclassified was calculated for each risk adjustment method using the classification based on the risk adjustment method as the correct classification. [b]Cohen's Kappa coefficient is a measure of agreement between unadjusted performance ratings and ratings based on the risk adjustment method with higher values indicating greater agreement

Table 38b: Agreement in groups: Case-mix adjusted STAT adherence scores

Groups Based on Case-mix Adjusted Estimates	Hierarchical Logistic Regression Model		
Logistic Regression Model	Bottom	Medium	Top
Bottom	7	1	0
Medium	1	26	0
Top	0	0	8
Percentage misclassified[a]	12.5%	3.7%	0%
Cohen's κ[b]	0.91		

[a]Percentage misclassified was calculated for each risk adjustment method using the classification based on hierarchical logistic regression model as the correct classification. [b]Cohen's Kappa coefficient is a measure of agreement between unadjusted performance ratings and ratings based on the risk adjustment method with higher values indicating greater agreement

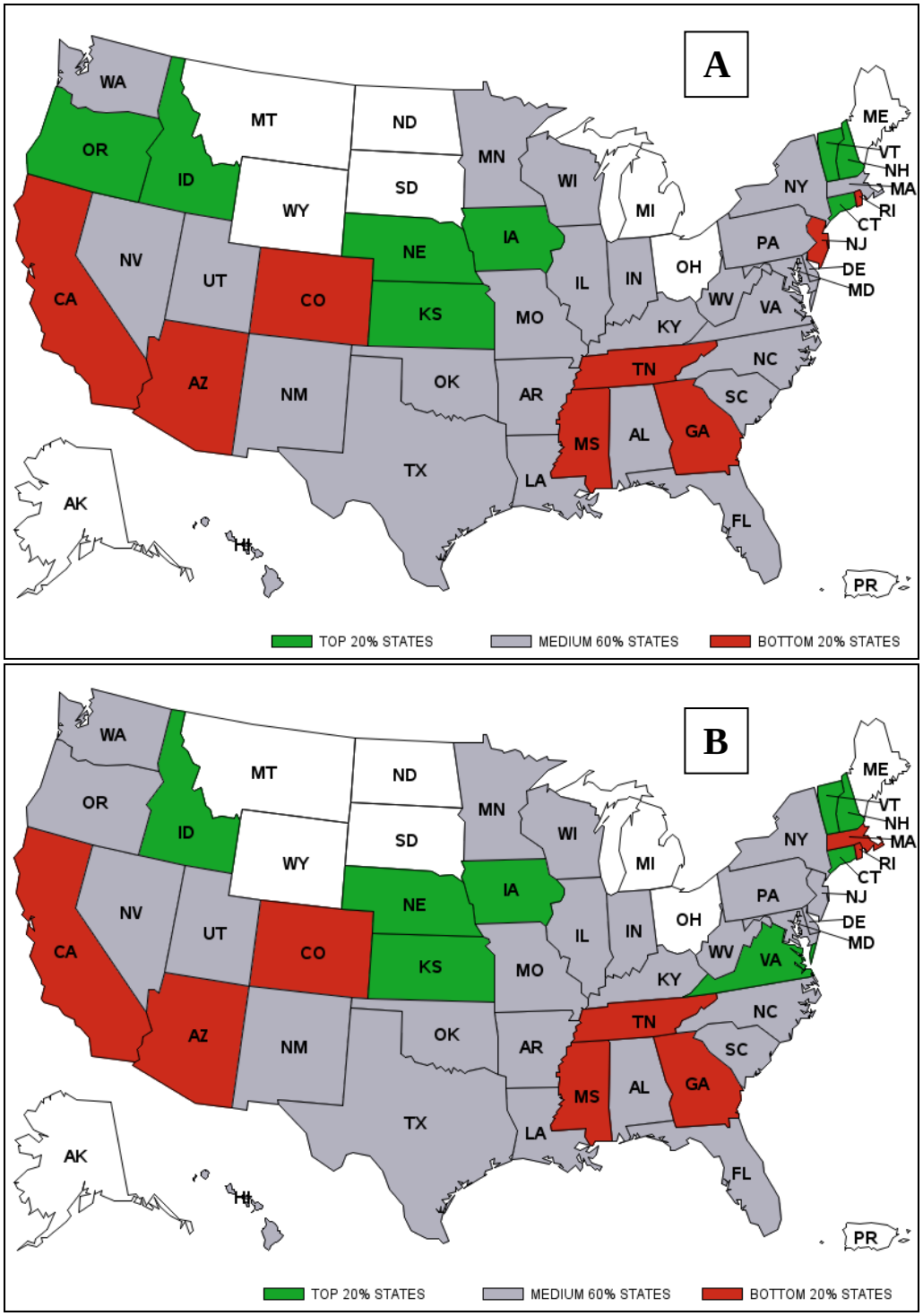

Figure 7. Interstate variations in STAT adherence measure: (A) Unadjusted scores, (B) Case-mix adjusted scores

Persistence with Antidepressant Medications (AD)

The AD measure was computed for 42 states and the District of Columbia. The number of patients eligible for the measure in each state and the proportion of patients with continuous use of antidepressant medications over an acute phase (12 weeks) and chronic phase (6 months) are shown in Table 39. The number of patients in the denominator ranged from 198 in District of Columbia to 12,401 in New York. The average proportion of patients with continuous use of antidepressants was 52.9% for acute phase measure and 31.3% for the chronic phase measure. The short term antidepressant use scores ranged from 44.7% in Mississippi to 61.9% in Vermont. Most states performed poorer than other measures on the chronic use of antidepressant measure, with scores ranging from 14.0% of Medicaid patients continuously using antidepressant medications over a period of six months in Hawaii to 40.6% in Vermont. States with scores above the national benchmark are highlighted in bold.

Table 39: Patients meeting the AD measure criteria by state

State	AD Measure	AD Acute	AD Chronic
	#patients in the denominator	%patients in the numerator	%patients in the numerator
Alabama	1318	50.76	33.08
Alaska	-	-	-
Arizona	3337	**57.12**	**34.85**
Arkansas	1032	50.68	28.59
California	8059	**53.96**	30.30
Colorado	1101	**61.31**	**37.87**
Connecticut	526	52.09	**35.55**
Delaware	489	**53.99**	**32.31**
District of Columbia	198	44.44	24.75
Florida	1641	52.35	**31.32**
Georgia	1645	48.57	25.05
Hawaii	314	40.76	14.01
Idaho	533	**58.54**	**34.90**
Illinois	3989	49.24	27.30
Indiana	2150	51.72	30.56

Table 39: Patients meeting the AD measure criteria by state (continued)

State	AD Measure	AD Acute	AD Chronic
	#patients in the denominator	%patients in the numerator	%patients in the numerator
Iowa	1481	**56.85**	**37.41**
Kansas	630	52.38	32.22
Kentucky	3112	52.41	30.17
Louisiana	1744	47.13	25.63
Maine	-	-	-
Maryland	1227	50.69	28.44
Massachusetts	2744	51.53	30.61
Michigan	-	-	-
Minnesota	3470	49.39	29.14
Mississippi	1239	44.71	21.87
Missouri	2689	52.77	**32.35**
Montana	-	-	-
Nebraska	555	**54.95**	**35.14**
Nevada	275	**61.45**	**37.82**
New Hampshire	497	**57.75**	**37.02**
New Jersey	1701	51.68	29.63
New Mexico	879	49.15	29.47
New York	12401	**54.64**	**33.10**
North Carolina	4499	**52.88**	**31.61**
North Dakota	-	-	-
Ohio	-	-	-
Oklahoma	1476	**56.78**	**35.09**
Oregon	1174	**55.71**	**32.28**
Pennsylvania	2112	**56.96**	**36.46**
Rhode Island	838	49.28	28.40
South Carolina	1624	49.57	26.35
South Dakota	-	-	-
Tennessee	5277	47.49	24.60
Texas	2462	**58.81**	**37.08**
Utah	780	**55.38**	**33.72**
Vermont	984	**61.89**	**40.55**
Virginia	1965	50.33	29.72
Washington	3422	**55.38**	**35.42**
West Virginia	1165	**54.94**	**35.11**
Wisconsin	5206	**52.94**	**32.17**
Wyoming	-	-	-
Total	93960	52.88	31.28

Demographic characteristics and co-morbidity measures for the eligible patient population are described in Table 40. The majority of the study population was comprised of women (67%) and the mean age (± SD) was 51 ± 10 years. A greater proportion of AD measure eligible population is female and White compared to the study population eligible for other measures. Non-persistent patients were slightly younger and a greater proportion of these patients were Black and Hispanic compared to persistent patients. The mean RxRisk scores were slightly higher in the persistent group compared to Medicaid beneficiaries who were non-persistent to antidepressants. Comorbidity burden measured using the CCI weighted and unweighted indices was lower in this population compared to Medicaid beneficiaries eligible for the PDC measures.

Table 40: Characteristics of Medicaid patients eligible for AD measure

Patient Characteristics	Prevalence(%) or Mean ± SD				
		Acute		Chronic	
	Overall	Persistent	Non-persistent	Persistent	Non-persistent
Age	37.37 ± 11.31	38.06 ± 11.28	36.59 ± 11.28	39.47 ± 11.19	36.41 ± 11.23
Sex					
• Female	81.46	80.66	82.36	79.12	82.53
• Male	18.54	19.34	17.64	20.88	17.47
Race					
• White	64.51	68.23	60.33	70.40	61.82
• Black	14.88	12.15	17.94	10.67	16.79
• Hispanic	12.35	11.46	13.36	10.69	13.11
• Other	8.26	8.16	8.37	8.24	8.27
RxRisk	6.17 ± 3.08	6.29 ± 3.15	6.03 ± 3.00	6.61 ± 3.22	5.97 ± 2.99
RxRisk (Weighted)	9.14 ± 5.15	9.24 ± 5.22	9.03 ± 5.07	9.62 ± 5.34	8.93 ± 5.05
CCI (Unweighted)	0.53 ± 0.94	0.56 ± 0.95	0.51 ± 0.93	0.62 ± 0.99	0.50 ± 0.91
CCI	0.82 ± 1.54	0.85 ± 1.56	0.78 ± 1.52	0.95 ± 1.62	0.75 ± 1.50

The results of the risk-adjusted analyses predicting persistence to antidepressants are presented in Table 41. The odds ratio estimates from the classical logistic regression model and the hierarchical logistic regression model with a random intercept are fairly similar. All parameters included in the models, including age, sex, race/ethnicity, and RxRisk score were found to be significant predictors of the AD measure. In case of the AD acute measure, c-statistic was 0.572 for both classical logistic and 0.571 for the hierarchical logistic regression models, showing weak discriminative ability for both models, compared to the other measures included in the study. However, the model performed better while predicting chronic phase persistence, with the c-statistic being 0.613 for both classical logistic and hierarchical logistic regression models, showing modest discriminative ability.

Race/ethnicity and gender were found to be strong predictors of antidepressant persistence in both models, with Blacks being 43% less likely to be persistent with the recommended antidepressant therapy over a 12-week time period compared to Whites in both regression models. Hispanics and other racial/ethnic groups were also less likely to be persistent compared to Whites. Age, sex and RxRisk were found to be significant predictors, but with modest association with persistence. Similar trends were observed in the model predicting chronic phase persistence. Blacks were 50% less likely to continue their medications over a 6-month period compared to Whites. Hispanic, other racial groups, and females were also found to be less likely to be persistent with antidepressants over a six month period compared to Whites and males, respectively, in both models.

Table 41: Odds ratio estimates of patient characteristics in the risk adjustment models

Baseline Characteristics	Classical Logistic Regression Model*	Hierarchical Logistic Regression Random Intercept Model*
	Point Estimate	Point Estimate
	AD Acute Measure	
Age	1.023	1.012
Sex		
• Female vs. Male	0.933	0.944
Race		
• Black vs. White	0.568	0.569
• Hispanic vs. White	0.718	0.649
• Other vs. White	0.812	0.774
RxRisk	1.018	1.021
	AD Chronic Measure	
Age	1.023	1.023
Sex		
• Female vs. Male	0.871	0.890
Race		
• Black vs. White	0.499	0.507
• Hispanic vs. White	0.643	0.587
• Other vs. White	0.777	0.732
RxRisk	1.047	1.051

*significant at $p < 0.001$

For the AD acute measure, in the hierarchical logistic regression model with a random-intercept, the state-level variance component was estimated to be 0.02262 (SE: 0.005893). Testing the null hypothesis of no random effects using a likelihood ratio test based on residual pseudo-likelihood yielded a chi-square of 307.94 ($p<0.0001$) indicating the presence of random effect. Therefore, case-mix adjusted scores based on hierarchical logistic regression model with a random intercept are presented in Table 42a. The residual intra class correlation coefficient (ρ) for the random intercept model was estimated to be 0.00978 which indicates that 0.98% of the

unexplained variation after controlling for patient level variables could be attributed to variation between states.

For the AD chronic measure, in the hierarchical logistic regression model with a random-intercept, the state-level variance component was estimated to be 0.04923 (SE: 0.01259). Testing the null hypothesis of no random effects using a likelihood ratio test based on residual pseudo-likelihood yielded a chi-square of 460.34 ($p<0.0001$) indicating the presence of a random effect. Therefore, case-mix adjusted scores based on hierarchical logistic regression model with a random intercept are presented in Table 42b. The residual intra class correlation coefficient (ρ) for the random intercept model was estimated to be 0.00978 which indicates that 0.98% of the unexplained variation after controlling for patient level variables could be attributed to variation between states.

Table 42a shows the agreement between the crude and case-mix adjusted scores. The case-mix adjusted AD acute measure scores ranged from 40.1% to 64.4%, a slight decrease in range compared to the crude estimates. Case-mix adjustment ranked the states differently with 14% (6) states ranked the same, 14% (6) states changing one rank and 72% (31) changing more than two positions. There was good agreement in rankings based on the crude scores and case-mix adjusted scores based on the hierarchical logistic regression model with a random intercept (Kendall's τ_b=0.71).

Table 42a: Agreement in ranks: Crude and case-mix adjusted AD Acute measure scores[a]

State	Unadjusted Estimates		Case-Mix Adjusted Estimates		Score difference	Rank difference
	Score (%)	Rank	Score (%)	Rank		
Alabama	50.76	**29**	50.78	**28**	-0.03	1
Alaska	-	**-**	-	**-**	-	-
Arizona	57.12	**7**	57.26	**5**	-0.15	2
Arkansas	50.68	**31**	49.92	**36**	0.75	-5
California	53.96	**18**	55.56	**9**	-1.60	9
Colorado	61.31	**3**	64.44	**1**	-3.13	2
Connecticut	52.09	**25**	52.77	**21**	-0.68	4
Delaware	53.99	**17**	55.01	**13**	-1.02	4
District of Columbia	44.44	**42**	50.43	**33**	-5.99	9
Florida	52.35	**24**	51.93	**24**	0.42	0
Georgia	48.57	**38**	48.92	**39**	-0.35	-1
Hawaii	40.76	**43**	40.08	**43**	0.69	0
Idaho	58.54	**5**	55.83	**8**	2.71	-3
Illinois	49.24	**36**	50.44	**32**	-1.20	4
Indiana	51.72	**26**	50.58	**31**	1.14	-5
Iowa	56.85	**9**	56.22	**7**	0.64	2
Kansas	52.38	**23**	51.83	**25**	0.55	-2
Kentucky	52.41	**22**	49.75	**37**	2.66	-15
Louisiana	47.13	**40**	48.80	**40**	-1.67	0
Maine	-	**-**	-	**-**	-	-
Maryland	50.69	**30**	53.29	**19**	-2.60	11
Massachusetts	51.53	**28**	50.59	**30**	0.94	-2
Michigan	-	**-**	-	**-**	-	-
Minnesota	49.39	**34**	50.36	**34**	-0.97	0
Mississippi	44.71	**41**	46.36	**42**	-1.64	-1
Missouri	52.77	**21**	51.02	**27**	1.75	-6
Montana	-	**-**	-	**-**	-	-
Nebraska	54.95	**14**	54.99	**14**	-0.04	0
Nevada	61.45	**2**	58.91	**4**	2.55	-2
New Hampshire	57.75	**6**	55.42	**10**	2.33	-4
New Jersey	51.68	**27**	52.13	**23**	-0.46	4
New Mexico	49.15	**37**	52.41	**22**	-3.26	15
New York	54.64	**16**	56.53	**6**	-1.89	10
North Carolina	52.88	**20**	53.02	**20**	-0.14	0
North Dakota	-	**-**	-	**-**	-	-

Table 42a: Agreement in ranks: Crude and case-mix adjusted AD Acute measure scores[a] (continued)

State	Unadjusted Estimates		Case-Mix Adjusted Estimates		Score difference	Rank difference
	Score (%)	Rank	Score (%)	Rank		
Ohio	-	-	-	-	-	-
Oklahoma	56.78	**10**	55.31	**11**	1.46	-1
Oregon	55.71	**11**	55.22	**12**	0.49	-1
Pennsylvania	56.96	**8**	54.84	**15**	2.12	-7
Rhode Island	49.28	**35**	51.04	**26**	-1.76	9
South Carolina	49.57	**33**	49.47	**38**	0.10	-5
South Dakota	-	-	-	-	-	-
Tennessee	47.49	**39**	46.95	**41**	0.54	-2
Texas	58.81	**4**	60.18	**2**	-1.36	2
Utah	55.38	**12**	53.65	**17**	1.73	-5
Vermont	61.89	**1**	60.14	**3**	1.75	-2
Virginia	50.33	**32**	50.71	**29**	-0.38	3
Washington	55.38	**13**	54.15	**16**	1.23	-3
West Virginia	54.94	**15**	50.00	**35**	4.93	-20
Wisconsin	52.94	**19**	53.34	**18**	-0.40	1
Wyoming	-	-	-	-	-	-
[b]Kendall's τ_b	0.71					

[a]AD acute measure scores could be generated for 43 states. Rankings ranged from 1-43

[b]Kendall's τ_b is a nonparametric measure of association based on the number of concordances and discordances in rankings based on unadjusted and risk-adjusted scores

Table 42b shows the agreement between the crude and case-mix adjusted scores. The case-mix adjusted AD chronic measure scores ranged from 13.4% to 42.5%, a decrease in range compared to the crude estimates. Case-mix adjustment ranked the states differently with 14% (6) states ranked the same, 21% (9) states changing one rank, and 65% (28) changing more than two positions. There was good agreement in rankings based on the crude scores and case-mix adjusted scores based on the hierarchical logistic regression model with a random intercept (Kendall's τ_b=0.71).

Table 42b: Agreement in ranks: Crude and case-mix adjusted AD Chronic measure scores[a]

State	Unadjusted Estimates		Case-Mix Adjusted Estimates		Score difference	Rank difference
	Score (%)	Rank	Score (%)	Rank		
Alabama	33.08	**17**	32.20	**20**	0.88	-3
Alaska	-	**-**	-	-	-	-
Arizona	34.85	**14**	35.10	**7**	-0.25	7
Arkansas	28.59	**33**	28.00	**36**	0.58	-3
California	30.30	**27**	31.86	**22**	-1.56	5
Colorado	37.87	**2**	42.49	**1**	-4.62	1
Connecticut	35.55	**8**	34.08	**14**	1.47	-6
Delaware	32.31	**19**	33.75	**15**	-1.44	4
District of Columbia	24.75	**40**	28.60	**35**	-3.85	5
Florida	31.32	**24**	30.10	**28**	1.23	-4
Georgia	25.05	**39**	24.51	**41**	0.53	-2
Hawaii	14.01	**43**	13.41	**43**	0.60	0
Idaho	34.90	**13**	33.23	**19**	1.67	-6
Illinois	27.30	**36**	29.32	**32**	-2.02	4
Indiana	30.56	**26**	29.99	**29**	0.57	-3
Iowa	37.41	**4**	38.03	**3**	-0.62	1
Kansas	32.22	**21**	32.10	**21**	0.12	0
Kentucky	30.17	**28**	27.49	**37**	2.69	-9
Louisiana	25.63	**38**	27.19	**38**	-1.56	0
Maine	-	**-**	-	-	-	-
Maryland	28.44	**34**	30.36	**27**	-1.92	7
Massachusetts	30.61	**25**	29.22	**33**	1.39	-8
Michigan	-	**-**	-	-	-	-
Minnesota	29.14	**32**	31.48	**24**	-2.35	8
Mississippi	21.87	**42**	23.28	**42**	-1.40	0
Missouri	32.35	**18**	30.73	**26**	1.62	-8
Montana	-	**-**	-	-	-	-
Nebraska	35.14	**10**	36.40	**5**	-1.26	5
Nevada	37.82	**3**	34.10	**13**	3.72	-10
New Hampshire	37.02	**6**	35.62	**6**	1.41	0
New Jersey	29.63	**30**	29.73	**31**	-0.10	-1
New Mexico	29.47	**31**	33.52	**16**	-4.05	15
New York	33.10	**16**	34.68	**9**	-1.57	7
North Carolina	31.61	**23**	31.50	**23**	0.11	0

Table 42b: Agreement in ranks: Crude and case-mix adjusted AD Chronic measure scores[a] (continued)

State	Unadjusted Estimates		Case-Mix Adjusted Estimates		Score difference	Rank difference
	Score (%)	Rank	Score (%)	Rank		
North Dakota	-	-	-	-	-	-
Ohio	-	-	-	-	-	-
Oklahoma	35.09	**12**	33.41	**17**	1.69	-5
Oregon	32.28	**20**	34.13	**12**	-1.85	8
Pennsylvania	36.46	**7**	35.09	**8**	1.37	-1
Rhode Island	28.40	**35**	31.10	**25**	-2.70	10
South Carolina	26.35	**37**	26.53	**39**	-0.18	-2
South Dakota	-	-	-	-	-	-
Tennessee	24.60	**41**	24.79	**40**	-0.19	1
Texas	37.08	**5**	37.50	**4**	-0.41	1
Utah	33.72	**15**	33.38	**18**	0.34	-3
Vermont	40.55	**1**	40.20	**2**	0.35	-1
Virginia	29.72	**29**	29.89	**30**	-0.17	-1
Washington	35.42	**9**	34.63	**10**	0.79	-1
West Virginia	35.11	**11**	28.77	**34**	6.34	-23
Wisconsin	32.17	**22**	34.19	**11**	-2.02	11
Wyoming	-	-	-	-	-	-
[b]Kendall's τ_b	0.71					

[a]AD chronic measure scores could be generated for 43 states. Rankings ranged from 1-43

[b]Kendall's τ_b is a nonparametric measure of association based on the number of concordances and discordances in rankings based on unadjusted and risk-adjusted scores

Additionally, states were classified into top (20%), medium and bottom (20%) performers based on the crude and case-mix adjusted scores. Results of the agreement in the grouping based on the unadjusted and risk-adjusted groupings are shown in Table 43a and Table 43b. Results based on the two risk adjustment models showed perfect agreement in classification of states into top, medium and bottom groups for the two AD measures (κ=1.00). There was modest agreement in classification of the states based on the crude and case-mix adjusted models when case-mix adjustment was conducted using both classical logistic regression (AD acute:

κ=0.57; AD chronic κ=0.65) and random intercept models (AD acute: κ=0.57; AD chronic κ=0.65).

An alternate methodology was proposed for classification of states by identifying outliers as low (high) quality outliers if the score for a state was significantly lower (or higher) than the average score according to the 95% CI of the measure. However, the distribution of scores for the 43 states is leptokurtic i.e., concentrated about the mean, therefore not conducive to identifying outliers. The results of agreement in classification based on this methodology are included in APPENDIX Table A-12. Choropleth maps depicting the top, medium and bottom performing states based on the crude and case-mix adjustment scores estimated from the hierarchical logistic regression model are depicted in Figures 8a and 8b. Additional choropleth maps illustrating the distribution of AD acute and AD chronic crude scores are presented in Figure A-10a and A-10b (APPENDIX).

Table 43a: Agreement in groups: Crude and case-mix adjusted AD measure scores

Groups Based on Crude Estimates	Groups Based on Case-mix Adjusted Estimates					
	Classical Logistic Regression Model			Hierarchical Logistic Regression Random Intercept Model		
AD Acute	Bottom	Medium	Top	Bottom	Medium	Top
Bottom (~20%)	5	3	0	5	3	0
Medium (60%)	3	22	2	3	22	2
Top (~20%)	0	2	6	0	2	6
Percentage misclassified[a]	37.5%	18.5%	25%	37.5%	18.5%	25%
Cohen's κ[b]	0.57			0.57		
AD Chronic	Bottom	Medium	Top	Bottom	Medium	Top
Bottom (~20%)	6	2	0	6	2	0
Medium (60%)	2	23	2	2	23	2
Top (~20%)	0	2	6	0	2	6
Percentage misclassified[a]	37.5%	14.8%	25%	37.5%	14.8%	25%
Cohen's κ[b]	0.65			0.65		

[a]Percentage misclassified was calculated for each risk adjustment method using the classification based on the risk adjustment method as the correct classification. [b]Cohen's Kappa coefficient is a measure of agreement between unadjusted performance ratings and ratings based on the risk adjustment method with higher values indicating greater agreement

Table 43b: Agreement in groups: Case-mix adjusted AD measure scores

Logistic Regression Model	Hierarchical Logistic Regression Model					
	AD Acute			AD Chronic		
AD Acute	Bottom	Medium	Top	Bottom	Medium	Top
Bottom (~20%)	8	0	0	8	0	0
Medium (60%)	0	27	0	0	27	0
Top (~20%)	0	0	8	0	0	8
Cohen's κ[a]	1.00			1.00		

[a]Cohen's Kappa coefficient is a measure of agreement between unadjusted performance ratings and ratings based on the risk adjustment method with higher values indicating greater agreement

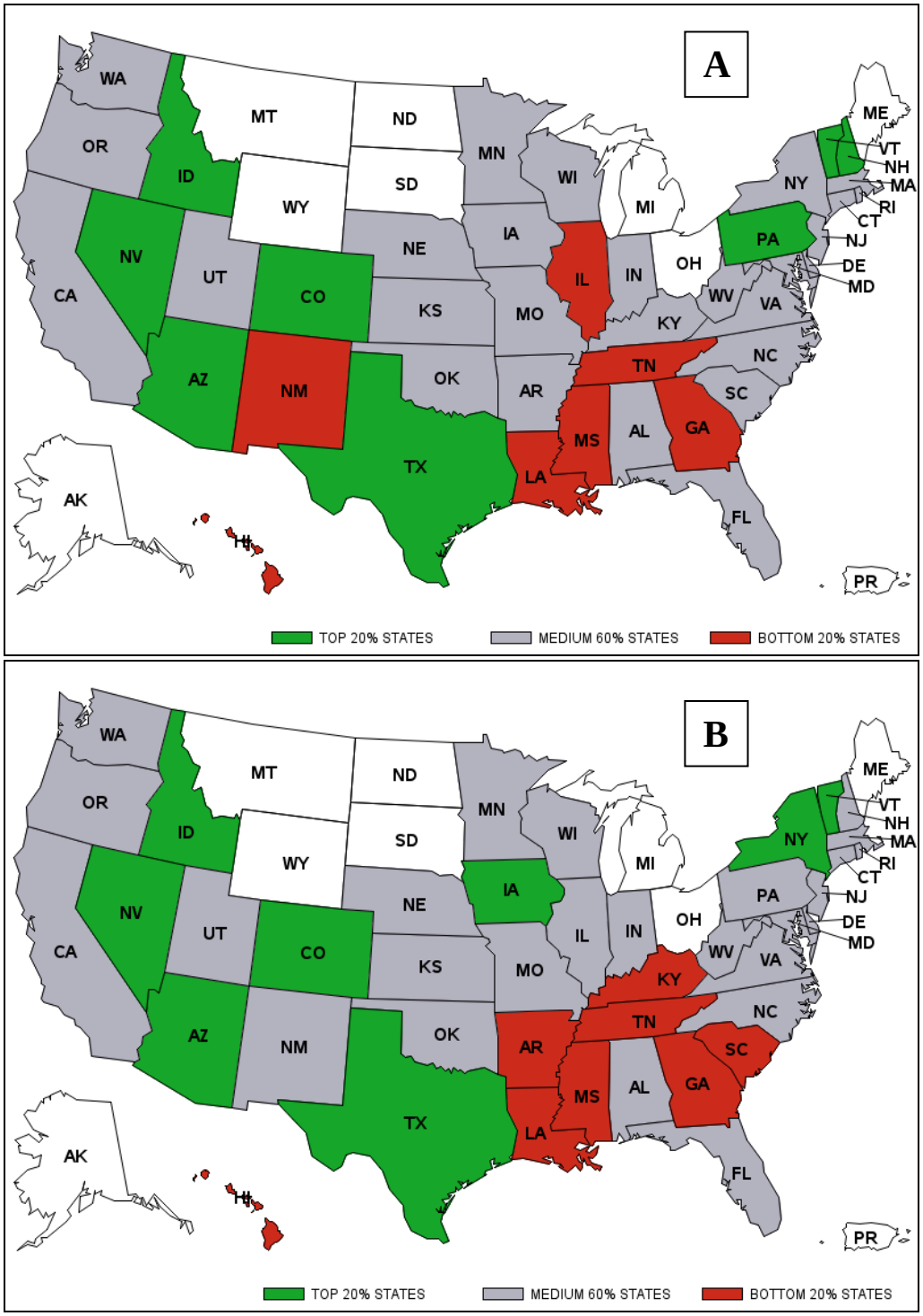

Figure 8a. Interstate variations in AD Acute measure: (A) Unadjusted scores, (B) Case-mix adjusted scores

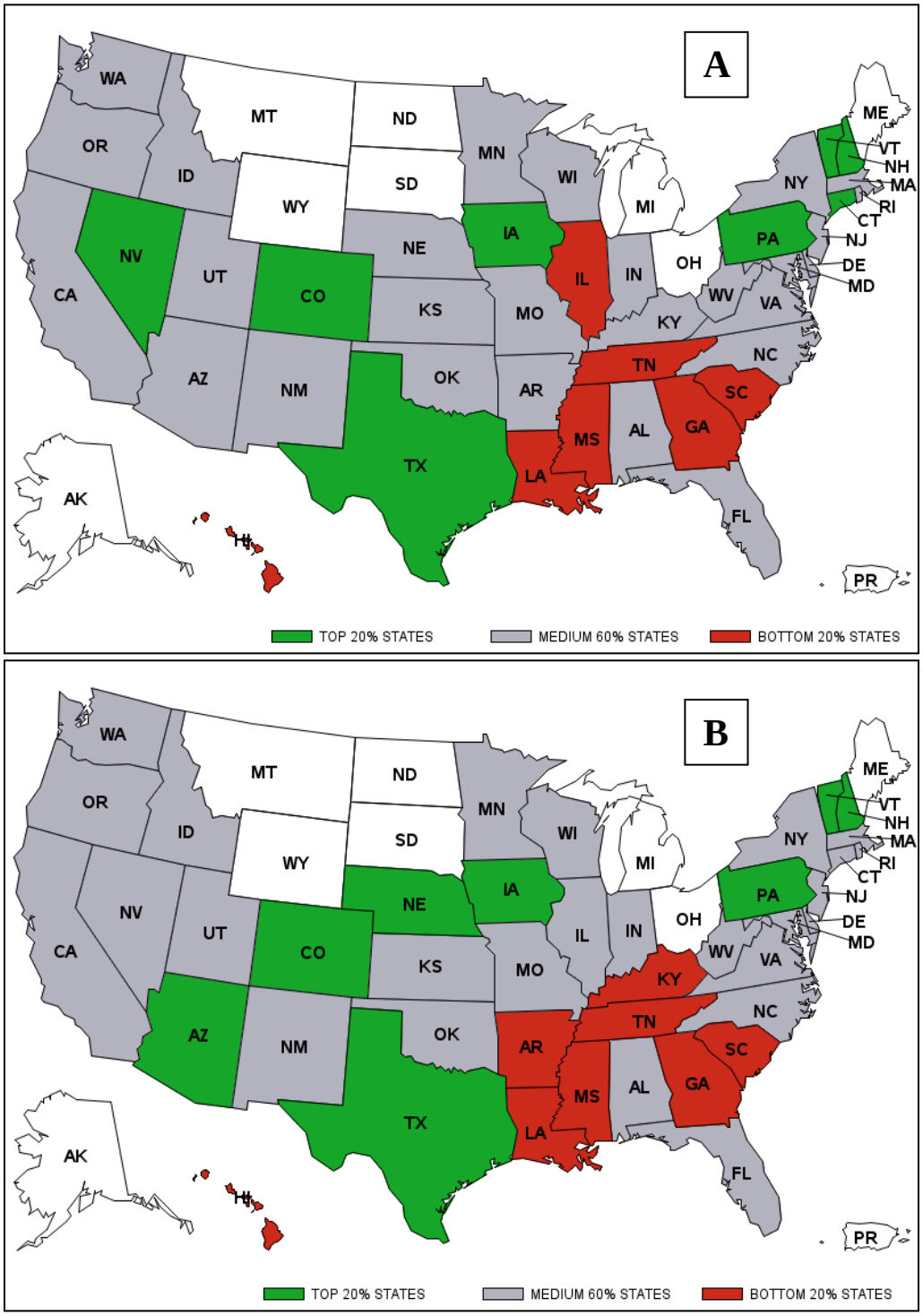

Figure 8b. Interstate variations in AD Chronic measure: (A) Unadjusted scores, (B) Case-mix adjusted scores

Statin use in coronary artery disease patients (CAD)

The CAD measure was computed for 42 states and the District of Columbia. The number of patients eligible for the measure in each state and the proportion of patients who filled at least one statin prescription are shown in Table 44. The number of eligible patients with a coronary artery disease diagnosis in 2007 ranged from 236 in Utah to 32,568 in New York. The average score on the CAD measure was 62.5% across all states, ranging from only 22.0% of Medicaid CAD patients being prescribed a statin in Oregon to 75.0% in Vermont. States with scores above the national benchmark are highlighted in bold.

Table 44: Patients meeting the CAD measure criteria by state

State	CAD	
	#patients in the denominator	%patients in the numerator
Alabama	3,448	**62.96**
Alaska	-	-
Arizona	3,652	60.51
Arkansas	1,548	52.97
California	16,534	61.82
Colorado	720	**63.89**
Connecticut	1,015	**72.51**
Delaware	581	**66.61**
District of Columbia	591	57.53
Florida	5,876	62.07
Georgia	4,335	61.96
Hawaii	721	49.79
Idaho	255	**68.24**
Illinois	5,890	**69.47**
Indiana	2,984	**64.81**
Iowa	850	**70.12**
Kansas	816	**63.24**
Kentucky	6,567	**70.03**
Louisiana	3,901	**66.55**
Maine	-	-
Maryland	2,751	55.47
Massachusetts	2,750	**66.98**
Michigan	-	-
Minnesota	988	**67.81**

Table 44: Patients meeting the CAD measure criteria by state (continued)

State	CAD	
	#patients in the denominator	%patients in the numerator
Mississippi	1,944	52.67
Missouri	3,369	**68.83**
Montana	-	-
Nebraska	329	**70.82**
Nevada	480	59.38
New Hampshire	275	**69.82**
New Jersey	3,718	59.01
New Mexico	783	**66.92**
New York	27,387	60.19
North Carolina	5,145	**68.59**
North Dakota	-	-
Ohio	-	-
Oklahoma	2,263	59.13
Oregon	817	22.03
Pennsylvania	2,275	**64.62**
Rhode Island	596	**69.80**
South Carolina	2,503	**64.44**
South Dakota	-	-
Tennessee	6,376	57.72
Texas	10,324	57.34
Utah	236	**64.83**
Vermont	372	**75.00**
Virginia	2,560	58.83
Washington	1,964	**72.76**
West Virginia	2,885	**71.58**
Wisconsin	2,017	**71.69**
Wyoming	-	-
Total	145,391	62.50

Demographic characteristics and co-morbidity measures for the eligible patient population are described in Table 45. The mean age (± SD) of the study population was 53 ± 9 years, and 55% were female. Medicaid beneficiaries with CAD who were prescribed a statin were slightly older and had more comorbidity burden than those that did not fill a statin prescription. A greater proportion of patients without a statin were Black compared to those with a statin medication.

Table 45: Characteristics of Medicaid patients eligible for CAD measure

Patient Characteristics	Prevalence(%) or Mean ± SD		
	Overall	Statin Rx	No Statin Rx
Age	52.54 ± 8.63	54.08 ± 7.19	49.97 ± 10.09
Sex			
• Female	55.05	53.43	57.74
• Male	44.95	46.57	42.26
Race			
• White	50.32	52.76	46.24
• Black	22.07	19.52	26.32
• Hispanic	13.99	13.65	14.57
• Other	13.62	14.06	12.87
RxRisk	7.76 ± 3.58	8.78 ± 3.15	6.06 ± 3.61
RxRisk (Weighted)	9.95 ± 5.97	10.48 ± 5.79	9.05 ± 6.16
CCI (Unweighted)	2.11 ± 1.63	2.19 ± 1.62	1.97 ± 1.64
CCI	3.02 ± 2.52	3.12 ± 2.43	2.86 ± 2.65

The results of the risk-adjusted analyses predicting statin use are presented in Table 46. The odds ratio estimates from the classical logistic regression model and the hierarchical logistic regression model with a random intercept are comparable. All parameters included in the models, including age, sex, race/ethnicity, and RxRisk score were found to be significant predictors of statin use. The c-statistic was 0.740 and 0.739 for the classical logistic and hierarchical logistic regression models, respectively, showing good discriminative ability for both models, compared to the other measures included in the study.

Race/ethnicity, gender and RxRisk score were found to be strong predictors of statin use in both models. Blacks were 24% less likely to fill a statin prescription than Whites. On the contrary, Hispanics and other racial groups were slightly more likely to use statins compared to Whites. Females were 36% less likely to use statins compared to males. Age was a significant predictor, but with modest association with statin use in both models.

Table 46: Odds ratio estimates of patient characteristics in the risk adjustment models

Baseline Characteristics	Classical Logistic Regression Model*	Hierarchical Logistic Regression Random Intercept Model*
	Point Estimate	Point Estimate
Age	1.045	1.046
Sex		
• Female vs. Male	0.633	0.644
Race		
• Black vs. White	0.767	0.758
• Hispanic vs. White	1.048	1.018
• Other vs. White	1.086	1.028
RxRisk	1.279	1.281

*significant at $p < 0.0001$

In the hierarchical logistic regression model with a random-intercept, the state-level variance component was estimated to be 0.06106 (SE: 0.01486). Testing the null hypothesis of no random effects using a likelihood ratio test based on residual pseudo-likelihood yielded a chi-square of 781.97 ($p<0.0001$) indicating the presence of a random effect. Therefore, case-mix adjusted scores based on hierarchical logistic regression model with a random intercept are presented in Table 47. The residual intra class correlation coefficient (ρ) for the random intercept model was estimated to be 0.01822 which indicates that 1.82% of the unexplained variation after controlling for patient level variables could be attributed to variation between states. State variable explained more variation in the CAD measure compared to the adherence and persistence measures studied.

Table 47 shows the agreement between the crude and case-mix adjusted scores. The case-mix adjusted CAD measure scores ranged from 35.47% to 73.49%, a decrease in range compared to the crude estimates. Case-mix adjustment ranked the states differently with 86% (37) changing more than two positions. There was modest agreement in rankings based on the

crude scores and case-mix adjusted scores based on the hierarchical logistic regression model with a random intercept (Kendall's τ_b=0.55).

Table 47: Agreement in ranks: Crude and case-mix adjusted CAD measure scores[a]

State	Unadjusted Estimates		Case-Mix Adjusted Estimates		Score difference	Rank difference
	Score (%)	Rank	Score (%)	Rank		
Alabama	62.96	**26**	60.06	**33**	2.91	-7
Alaska	-	**-**	-	**-**	-	-
Arizona	60.51	**30**	62.84	**25**	-2.33	5
Arkansas	52.97	**40**	59.22	**35**	-6.24	5
California	61.82	**29**	65.14	**15**	-3.31	14
Colorado	63.89	**24**	63.14	**23**	0.75	1
Connecticut	72.51	**3**	67.01	**8**	5.50	-5
Delaware	66.61	**18**	68.89	**5**	-2.28	13
District of Columbia	57.53	**37**	62.13	**29**	-4.60	8
Florida	62.07	**27**	57.94	**41**	4.13	-14
Georgia	61.96	**28**	60.16	**32**	1.80	-4
Hawaii	49.79	**42**	58.49	**40**	-8.69	2
Idaho	68.24	**14**	63.04	**24**	5.20	-10
Illinois	69.47	**11**	69.45	**3**	0.03	8
Indiana	64.81	**21**	62.33	**28**	2.48	-7
Iowa	70.12	**7**	66.46	**12**	3.66	-5
Kansas	63.24	**25**	60.00	**34**	3.23	-9
Kentucky	70.03	**8**	63.33	**21**	6.71	-13
Louisiana	66.55	**19**	64.13	**19**	2.42	0
Maine	-	**-**	-	**-**	-	-
Maryland	55.47	**39**	58.86	**37**	-3.39	2
Massachusetts	66.98	**16**	67.46	**6**	-0.48	10
Michigan	-	**-**	-	**-**	-	-
Minnesota	67.81	**15**	66.67	**9**	1.15	6
Mississippi	52.67	**41**	60.54	**31**	-7.86	10
Missouri	68.83	**12**	66.48	**11**	2.36	1
Montana	-	**-**	-	**-**	-	-
Nebraska	70.82	**6**	64.41	**17**	6.41	-11
Nevada	59.38	**32**	55.82	**42**	3.56	-10
New Hampshire	69.82	**9**	64.69	**16**	5.13	-7
New Jersey	59.01	**34**	62.79	**26**	-3.78	8

Table 47: Agreement in ranks: Crude and case-mix adjusted CAD measure scores[a] (continued)

State	Unadjusted Estimates		Case-Mix Adjusted Estimates		Score difference	Rank difference
	Score (%)	Rank	Score (%)	Rank		
New Mexico	66.92	**17**	62.33	**27**	4.59	-10
New York	60.19	**31**	67.04	**7**	-6.85	24
North Carolina	68.59	**13**	64.25	**18**	4.34	-5
North Dakota	-	**-**	-	**-**	-	-
Ohio	-	**-**	-	**-**	-	-
Oklahoma	59.13	**33**	58.70	**39**	0.43	-6
Oregon	22.03	**43**	35.47	**43**	-13.44	0
Pennsylvania	64.62	**22**	64.07	**20**	0.55	2
Rhode Island	69.80	**10**	73.38	**2**	-3.58	8
South Carolina	64.44	**23**	66.59	**10**	-2.15	13
South Dakota	-	**-**	-	**-**	-	-
Tennessee	57.72	**36**	61.68	**30**	-3.97	6
Texas	57.34	**38**	59.17	**36**	-1.83	2
Utah	64.83	**20**	65.81	**14**	-0.98	6
Vermont	75.00	**1**	73.49	**1**	1.51	0
Virginia	58.83	**35**	58.76	**38**	0.07	-3
Washington	72.76	**2**	65.90	**13**	6.86	-11
West Virginia	71.58	**5**	63.22	**22**	8.36	-17
Wisconsin	71.69	**4**	69.28	**4**	2.41	0
Wyoming	-	**-**	-	**-**	-	-
[b]Kendall's τ_b	0.55					

[a]CAD measure scores could be generated for 43 states. Rankings ranged from 1-43

[b]Kendall's τ_b is a nonparametric measure of association based on the number of concordances and discordances in rankings based on unadjusted and risk-adjusted scores

Additionally, states were classified into top (20%), medium and bottom (20%) performers based on the crude and case-mix adjusted scores. Results of the agreement in the grouping based on the unadjusted and risk-adjusted groupings are shown in Table 48a and Table 48b. Results based on the two risk adjustment models showed perfect agreement in classification of states into top, medium and bottom groups (κ=1.00). There was poor agreement in classification of the states based on the crude and case-mix adjusted models when case-mix

adjustment was conducted using both classical logistic regression (κ=0.22) and random intercept models (κ=0.22).

An alternate methodology was proposed for classification of states by identifying outliers as low (high) quality outliers if the score for a state was significantly lower (or higher) than the average score according to the 95% CI of the measure. However, the distribution of scores for the 43 states is leptokurtic i.e., concentrated about the mean, therefore not conducive to identifying outliers. The results of agreement in classification based on this methodology are included in APPENDIX Table A-13. Choropleth maps depicting the top, medium and bottom performing states based on the crude and case-mix adjustment scores estimated from the hierarchical logistic regression model are depicted in Figure 9. Additional choropleth maps illustrating the distribution of crude adherence scores are presented in APPENDIX Figure A-11.

Table 48a: Agreement in groups: Crude and case-mix adjusted CAD measure scores

Groups Based on Crude Estimates	Groups Based on Case-mix Adjusted Estimates					
	Classical Logistic Regression Model			Hierarchical Logistic Regression Random Intercept Model		
	Bottom	Medium	Top	Bottom	Medium	Top
Bottom (~20%)	4	4	0	4	4	0
Medium (60%)	4	18	5	4	18	5
Top (~20%)	0	5	3	0	5	3
Percentage misclassified[a]	50.0%	33.3%	62.5%	50.0%	33.3%	62.5%
Cohen's κ[b]	0.22			0.22		

[a]Percentage misclassified was calculated for each risk adjustment method using the classification based on the risk adjustment method as the correct classification. [b]Cohen's Kappa coefficient is a measure of agreement between unadjusted performance ratings and ratings based on the risk adjustment method with higher values indicating greater agreement

Table 48b: Agreement in groups: Case-mix adjusted CAD measure scores

Groups Based on Case-mix Adjusted Estimates	Hierarchical Logistic Regression Model		
Logistic Regression Model	Bottom	Medium	Top
Bottom	8	0	0
Medium	0	27	0
Top	0	0	8
Percentage misclassified[a]	0%	0%	0%
Cohen's κ[b]	1.00		

[a]Percentage misclassified was calculated for each risk adjustment method using the classification based on hierarchical logistic regression model as the correct classification. [b]Cohen's Kappa coefficient is a measure of agreement between unadjusted performance ratings and ratings based on the risk adjustment method with higher values indicating greater agreement

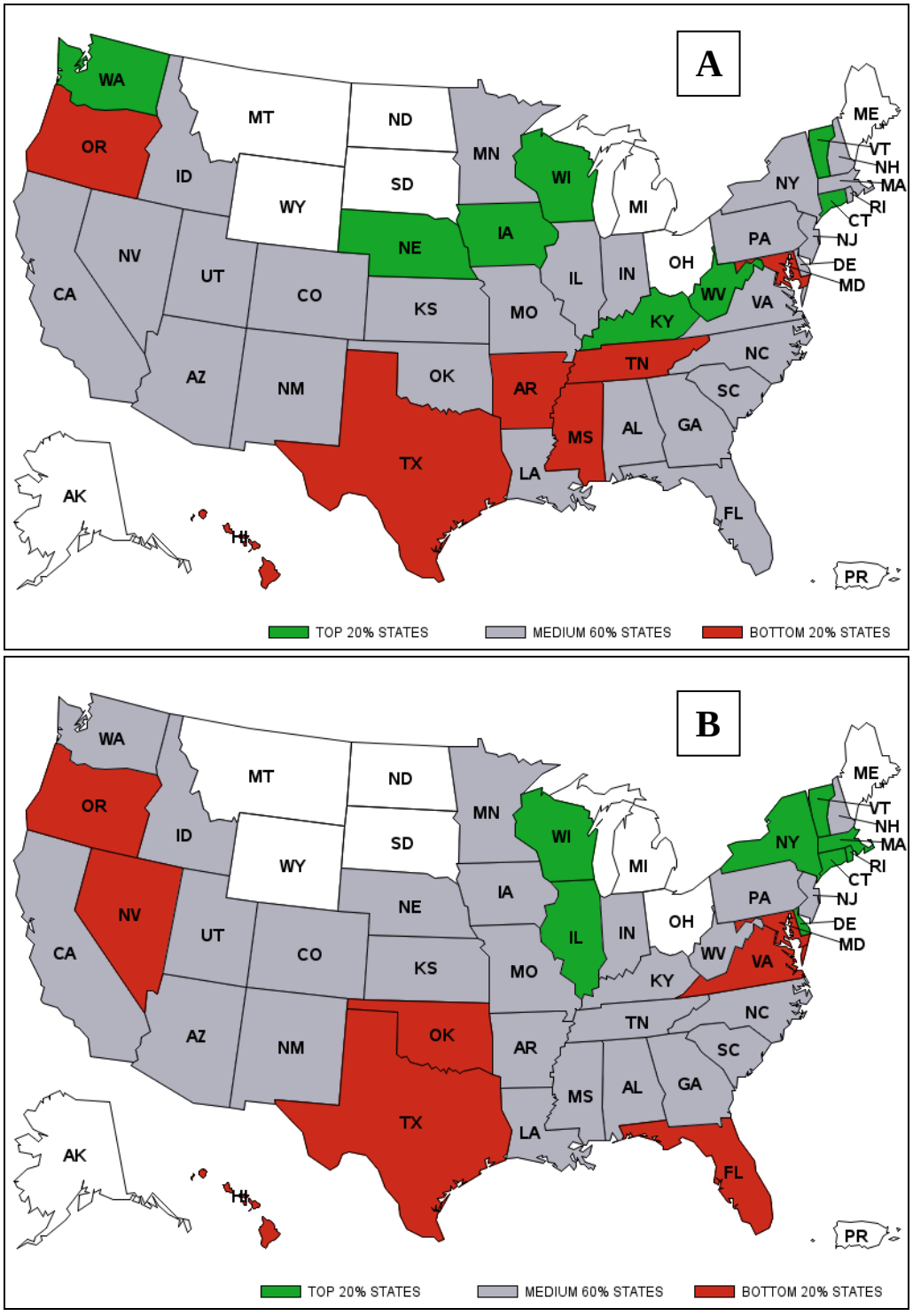

Figure 9. Interstate variations in CAD measure:
(A) Unadjusted scores, (B) Case-mix adjusted scores

Inhaled corticosteroid use in asthma patients (ICS)

The ICS measure was computed for 42 states and the District of Columbia. The number of patients eligible for the measure in each state and the proportion of patients who filled an inhaled corticosteroid or similar medication are shown in Table 49. The number of patients in the denominator ranged from 608 in District of Columbia to 50,517 in New York. The average ICS score across all states was 68.1%, ranging from 57.0% of Medicaid patients filling an inhaled corticosteroid or similar medication prescription in Tennessee to 75.4% in Delaware. States with scores above the national benchmark are highlighted in bold.

Table 49: Patients meeting the ICS measure criteria by state

State	ICS	
	#patients in the denominator	%patients in the numerator
Alabama	4,523	**69.18**
Alaska	-	-
Arizona	7,279	62.01
Arkansas	2,414	62.39
California	30,512	**70.71**
Colorado	2,000	61.35
Connecticut	1,658	**74.61**
Delaware	1,536	**75.39**
District of Columbia	608	67.60
Florida	8,043	**70.87**
Georgia	4,071	62.74
Hawaii	939	65.71
Idaho	878	61.50
Illinois	15,308	**70.73**
Indiana	4,511	59.17
Iowa	2,825	67.26
Kansas	1,597	64.50
Kentucky	8,267	61.32
Louisiana	4,305	**68.97**
Maine	-	-
Maryland	4,795	65.07
Massachusetts	7,716	62.69
Michigan	-	-
Minnesota	4,970	66.88

Table 49: Patients meeting the ICS measure criteria by state (continued)

State	ICS	
	#patients in the denominator	%patients in the numerator
Mississippi	2,333	**70.90**
Missouri	6,701	65.87
Montana	-	-
Nebraska	1,243	**73.61**
Nevada	748	64.71
New Hampshire	936	67.52
New Jersey	7,261	**73.43**
New Mexico	2,388	61.89
New York	50,517	**72.50**
North Carolina	9,268	**70.30**
North Dakota	-	-
Ohio	-	-
Oklahoma	3,174	65.00
Oregon	887	65.73
Pennsylvania	4,588	**69.31**
Rhode Island	1,812	**70.42**
South Carolina	3,098	66.24
South Dakota	-	-
Tennessee	12,424	56.99
Texas	8,411	**74.71**
Utah	1,146	66.58
Vermont	1,628	**70.52**
Virginia	5,103	67.88
Washington	7,018	60.06
West Virginia	3,575	67.22
Wisconsin	9,029	66.90
Wyoming	-	-
Total	262,043	68.12

Demographic characteristics and co-morbidity measures for the eligible patient population are described in Table 50. The majority of the study population was comprised of women (76%) and the mean age (± SD) was 38 ± 12 years, younger than the population eligible for other measures being studied. Approximately 50% of the study population was White, with a greater proportion of Blacks and Hispanics being prescribed an inhaled corticosteroid or similar medications compared to Whites. This trend was in contrary to the race/ethnicity effect seen in

case of other measures being studied. Age, gender and comorbidity burden of patients included in the numerator for the ICS measure were similar to those not included in the numerator.

Table 50: Characteristics of Medicaid patients eligible for ICS measure

Patient Characteristics	Prevalence(%) or Mean ± SD		
	Overall	**ICS Rx**	**No ICS Rx**
Age	37.55 ± 11.62	37.63 ± 11.68	37.36 ± 11.49
Sex			
• Female	76.02	76.66	74.65
• Male	23.98	23.34	25.35
Race			
• White	50.59	48.87	54.27
• Black	24.74	25.04	24.10
• Hispanic	15.46	16.38	13.50
• Other	9.21	9.71	8.14
RxRisk	5.55 ± 3.04	5.63 ± 3.10	5.39 ± 2.91
RxRisk (Weighted)	10.21 ± 4.46	10.25 ± 4.56	10.13 ± 4.23
CCI (Unweighted)	0.98 ± 0.99	1.05 ± 0.98	0.84 ± 0.98
CCI	1.30 ± 1.56	1.36 ± 1.55	1.17 ± 1.58

*ICS – Inhaled corticosteroid or similar medications

The results of the risk-adjusted analyses predicting inhaled corticosteroid medication use are presented in Table 51. All parameters included in the models, including age, sex, race/ethnicity, and RxRisk score were found to be significant predictors at P<0.0001 level. However, the c-statistic was 0.545 and 0.541 for the classical logistic and hierarchical logistic regression models, showing weak discriminative ability for both models. Therefore, crude estimates could be used to describe the ICS measure. Ranks and groups based on case-mix adjusted scores are expected to be similar to those based on crude estimates for the ICS measure. However, for uniformity in presentation of data, the case-mix adjusted estimates will be described for this measure.

The odds ratio estimates for age and RxRisk variables from the classical logistic regression model and the hierarchical logistic regression model with a random intercept were similar. The effect of gender on ICS use was different in the two models, with classical logistic regression estimate indicating that females were 9% more likely to use inhaled corticosteroid or similar medications to manage their persistent asthma compared to males. Contrarily, females were found to be 11% less likely to use ICS than males using the hierarchical logistic regression model with a random intercept. Black, Hispanic and other racial groups were observed to be more likely to use ICS or similar medications compared to white in the two models.

Table 51: Odds ratio estimates of patient characteristics in the risk adjustment models

Baseline Characteristics	Classical Logistic Regression Model*	Hierarchical Logistic Regression Random Intercept Model*
	Point Estimate	Point Estimate
Age	0.998	0.997
Sex		
• Female vs. Male	1.091	0.890
Race		
• Black vs. White	1.175	1.070
• Hispanic vs. White	1.381	1.178
• Other vs. White	1.354	1.311
RxRisk	1.032	1.036

*significant at $p < 0.0001$

In the hierarchical logistic regression model with a random-intercept, the state-level variance component was estimated to be 0.04055 (SE: 0.009211). Testing the null hypothesis of no random effects using a likelihood ratio test based on residual pseudo-likelihood yielded a chi-square of 2152.62 ($p<0.0001$) indicating the presence of random effect. Therefore, case-mix adjusted scores based on hierarchical logistic regression model with a random intercept are presented in Table 52. The residual intra class correlation coefficient (ρ) for the random intercept

model was estimated to be 0.01218 which indicates that 1.22% of the unexplained variation after controlling for patient level variables could be attributed to variation between states.

Table 52 shows the agreement between the crude and case-mix adjusted scores. The case-mix adjusted ICS measure scores ranged from 40.46% to 65.43%, a decrease in range compared to the crude estimates. Case-mix adjustment ranked the states differently with 32% (14) states ranked the same and 37% (16) changing more than two positions. There was close to perfect agreement in rankings based on the crude scores and case-mix adjusted scores based on the hierarchical logistic regression model with a random intercept (Kendall's τ_b=0.93).

Table 52: Agreement in ranks: Crude and case-mix adjusted ICS measure scores[a]

State	Unadjusted Estimates		Case-Mix Adjusted Estimates		Score difference	Rank difference
	Score (%)	Rank	Score (%)	Rank		
Alabama	69.18	**15**	68.70	**15**	0.48	0
Alaska	-	**-**	-	**-**	-	-
Arizona	62.01	**36**	62.10	**36**	-0.09	0
Arkansas	62.39	**35**	63.11	**33**	-0.73	2
California	70.71	**10**	70.38	**10**	0.33	0
Colorado	61.35	**39**	59.37	**41**	1.98	-2
Connecticut	74.61	**3**	73.11	**4**	1.49	-1
Delaware	75.39	**1**	75.57	**1**	-0.18	0
District of Columbia	67.60	**18**	67.15	**22**	0.45	-4
Florida	70.87	**8**	69.75	**13**	1.12	-5
Georgia	62.74	**33**	62.30	**35**	0.44	-2
Hawaii	65.71	**28**	64.10	**32**	1.60	-4
Idaho	61.50	**38**	61.98	**37**	-0.48	1
Illinois	70.73	**9**	70.80	**9**	-0.06	0
Indiana	59.17	**42**	59.33	**42**	-0.17	0
Iowa	67.26	**20**	67.13	**23**	0.13	-3
Kansas	64.50	**32**	64.58	**31**	-0.08	1
Kentucky	61.32	**40**	61.04	**38**	0.28	2
Louisiana	68.97	**16**	68.50	**16**	0.47	0
Maine	-	**-**	-	**-**	-	-
Maryland	65.07	**29**	65.36	**28**	-0.29	1

Table 52: Agreement in ranks: Crude and case-mix adjusted ICS measure scores[a] (continued)

State	Unadjusted Estimates		Case-Mix Adjusted Estimates		Score difference	Rank difference
	Score (%)	Rank	Score (%)	Rank		
Massachusetts	62.69	**34**	63.04	**34**	-0.35	0
Michigan	-	**-**	-	**-**	-	-
Minnesota	66.88	**23**	67.20	**21**	-0.32	2
Mississippi	70.90	**7**	71.31	**7**	-0.41	0
Missouri	65.87	**26**	66.09	**25**	-0.22	1
Montana	-	**-**	-	**-**	-	-
Nebraska	73.61	**4**	73.75	**3**	-0.14	1
Nevada	64.71	**31**	64.62	**30**	0.08	1
New Hampshire	67.52	**19**	68.12	**17**	-0.60	2
New Jersey	73.43	**5**	73.06	**5**	0.37	0
New Mexico	61.89	**37**	60.60	**39**	1.29	-2
New York	72.50	**6**	72.03	**6**	0.47	0
North Carolina	70.30	**13**	69.68	**14**	0.61	-1
North Dakota	-	**-**	-	**-**	-	-
Ohio	-	**-**	-	**-**	-	-
Oklahoma	65.00	**30**	65.23	**29**	-0.23	1
Oregon	65.73	**27**	65.95	**27**	-0.22	0
Pennsylvania	69.31	**14**	70.05	**12**	-0.74	2
Rhode Island	70.42	**12**	70.18	**11**	0.24	1
South Carolina	66.24	**25**	66.35	**24**	-0.11	1
South Dakota	-	**-**	-	**-**	-	-
Tennessee	56.99	**43**	57.55	**43**	-0.55	0
Texas	74.71	**2**	74.14	**2**	0.57	0
Utah	66.58	**24**	67.23	**20**	-0.65	4
Vermont	70.52	**11**	71.18	**8**	-0.66	3
Virginia	67.88	**17**	67.81	**18**	0.07	-1
Washington	60.06	**41**	59.99	**40**	0.07	1
West Virginia	67.22	**21**	67.49	**19**	-0.28	2
Wisconsin	66.90	**22**	65.97	**26**	0.92	-4
Wyoming	-	**-**	-	**-**	-	-
[b]Kendall's τ_b	0.93					

[a]ICS measure scores could be generated for 43 states. Rankings ranged from 1-43

[b]Kendall's τ_b is a nonparametric measure of association based on the number of concordances and discordances in rankings based on unadjusted and risk-adjusted scores

Additionally, states were classified into top (20%), medium and bottom (20%) performers based on the crude and case-mix adjusted scores. Results of the agreement in the grouping based on the unadjusted and risk-adjusted groupings are shown in Table 53a and Table 53b. Results based on the two risk adjustment models showed close to perfect agreement in classification of states into top, medium and bottom groups (κ=0.91). There was good agreement in classification of the states based on the crude and case-mix adjusted models when case-mix adjustment was conducted using both classical logistic regression (κ=0.83) and random intercept models (κ=0.91).

An alternate methodology was proposed for classification of states by identifying outliers as low (high) quality outliers if the score for a state was significantly lower (or higher) than the average score according to the 95% CI of the measure. However, the distribution of scores for the 43 states is leptokurtic i.e., concentrated about the mean, therefore not conducive to identifying outliers. The results of agreement in classification based on this methodology are included in Table A-14 (APPENDIX). Choropleth maps depicting the top, medium and bottom performing states based on the crude and case-mix adjustment scores estimated from the hierarchical logistic regression model are depicted in Figure 10. Additional choropleth maps illustrating the distribution of crude adherence scores are presented in Figure A-12 (APPENDIX).

Table 53a: Agreement in groups: Crude and case-mix adjusted ICS measure scores

Groups Based on Crude Estimates	Groups Based on Case-mix Adjusted Estimates					
	Classical Logistic Regression Model			Hierarchical Logistic Regression Random Intercept Model		
	Bottom	Medium	Top	Bottom	Medium	Top
Bottom (~20%)	7	1	0	8	0	0
Medium (60%)	1	25	1	0	26	1
Top (~20%)	0	1	7	0	1	7
Percentage misclassified[a]	12.5%	7.4%	12.5%	0%	3.7%	12.5%
Cohen's κ[b]	0.83			0.91		

[a]Percentage misclassified was calculated for each risk adjustment method using the classification based on the risk adjustment method as the correct classification. [b]Cohen's Kappa coefficient is a measure of agreement between unadjusted performance ratings and ratings based on the risk adjustment method with higher values indicating greater agreement

Table 53b: Agreement in groups: Case-mix adjusted ICS measure scores

Groups Based on Case-mix Adjusted Estimates	Hierarchical Logistic Regression Model		
Logistic Regression Model	Bottom	Medium	Top
Bottom	7	1	0
Medium	1	26	0
Top	0	0	8
Percentage misclassified[a]	12.5%	3.7%	0%
Cohen's κ[b]	0.91		

[a]Percentage misclassified was calculated for each risk adjustment method using the classification based on hierarchical logistic regression model as the correct classification. [b]Cohen's Kappa coefficient is a measure of agreement between unadjusted performance ratings and ratings based on the risk adjustment method with higher values indicating greater agreement

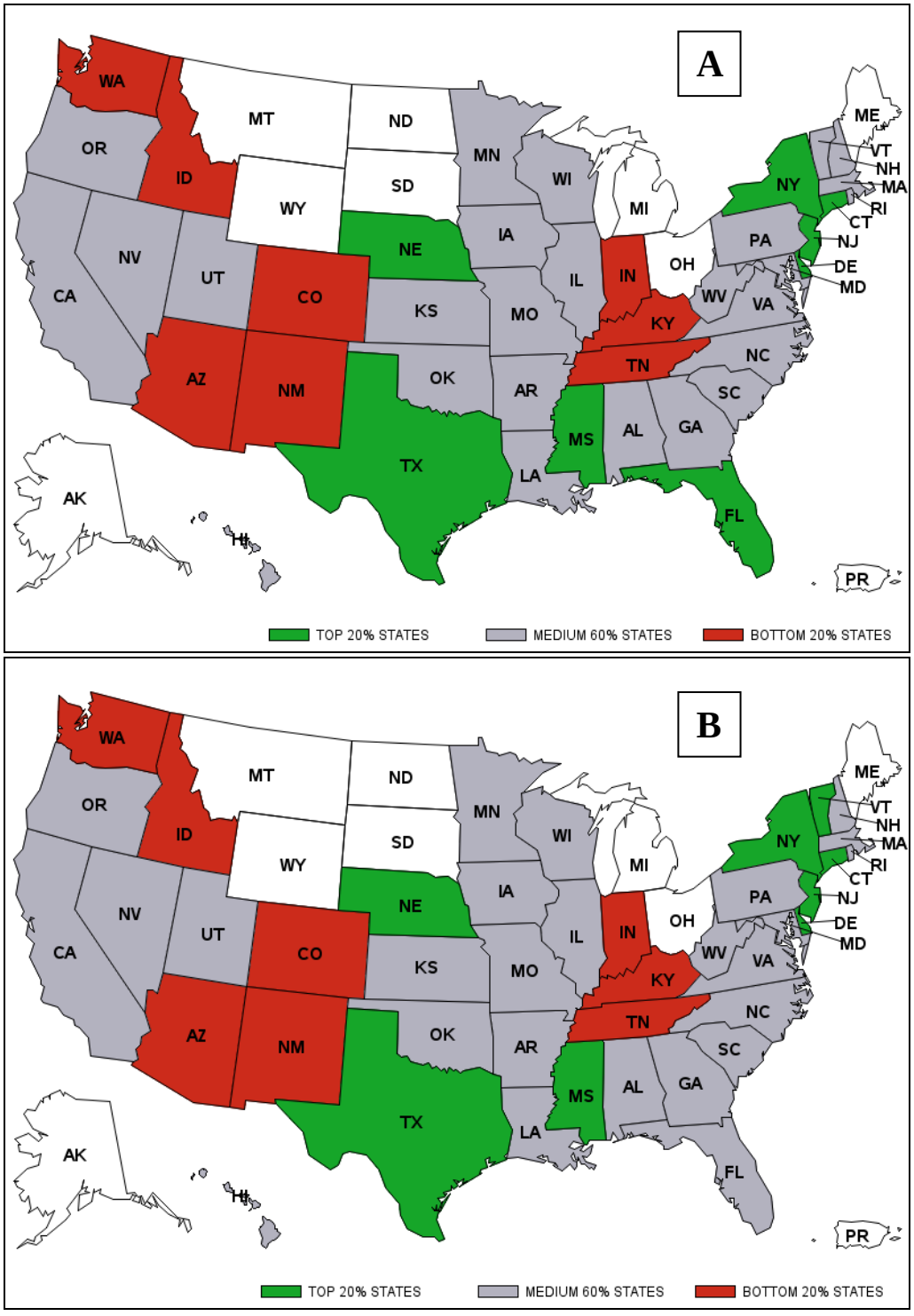

Figure 10. Interstate variations in ICS measure:
(A) Unadjusted scores, (B) Case-mix adjusted scores

Beta-blocker use in myocardial infarction patients (MI1)

The MI1 measure assessing the proportion of patients with an acute myocardial infarction who filled a beta-blocker prescription within 30 days of discharge from the hospital could be computed for 35 states and the District of Columbia. The number of patients eligible for the measure in each state and the proportion of patients who filled a beta-blocker prescription are shown in Table 54. The number of patients in the denominator ranged from 33 patients in New Mexico to 967 in California. The average MI1 measure score across all states was 49.0% across all states, ranging from 19.7% in Oregon to 66.7% in Wisconsin. States with scores above the national benchmark are highlighted in bold.

Table 54: Patients meeting the MI1 measure criteria by state

State	MI1	
	#patients in the denominator	%patients in the numerator
Alabama	184	**63.04**
Alaska	-	-
Arizona	281	**58.72**
Arkansas	91	40.66
California	967	47.57
Colorado	54	**62.96**
Connecticut	57	**54.39**
Delaware	-	-
District of Columbia	55	**49.09**
Florida	414	**51.45**
Georgia	276	**57.97**
Hawaii	71	**54.93**
Idaho	-	-
Illinois	474	**50.63**
Indiana	214	38.79
Iowa	82	26.83
Kansas	55	45.45
Kentucky	325	**65.23**
Louisiana	246	**57.72**
Maine	-	-
Maryland	134	42.54
Massachusetts	180	**68.33**

Table 54: Patients meeting the MI1 measure criteria by state (continued)

State	MI1	
	#patients in the denominator	%patients in the numerator
Michigan	-	-
Minnesota	88	**59.09**
Mississippi	137	37.96
Missouri	243	**50.21**
Montana	-	-
Nebraska	-	-
Nevada	34	32.35
New Hampshire	-	-
New Jersey	141	**58.87**
New Mexico	33	**60.61**
New York	826	32.32
North Carolina	430	**53.26**
North Dakota	-	-
Ohio	-	-
Oklahoma	149	44.97
Oregon	71	19.72
Pennsylvania	212	42.92
Rhode Island	-	-
South Carolina	182	**50.55**
South Dakota	-	-
Tennessee	289	**56.40**
Texas	667	34.03
Utah	-	-
Vermont	-	-
Virginia	133	**52.63**
Washington	154	**61.69**
West Virginia	200	**63.50**
Wisconsin	144	**66.67**
Wyoming	-	-
Total	8,293	49.01

Demographic characteristics and co-morbidity measures for the eligible patient population are described in Table 55. The majority of the study population was comprised of Whites (52%) and the mean age (± SD) was 53 ± 8 years. A greater proportion of the study population is males compared to the other medication-related measures included in the study. A greater proportion of the patients who did not fill a beta-blocker prescription were females, and

Hispanic compared to those who filled the beta-blocker prescription post discharge. The mean RxRisk scores were slightly higher in the MI patients who filled a beta-blocker prescription. CCI weighted and unweighted indices were similar across the two groups.

Table 55: Characteristics of Medicaid patients eligible for MI1 measure

Patient Characteristics	Prevalence(%) or Mean ± SD		
	Overall	**Beta-blocker Rx**	**No Beta-blocker Rx**
Age	52.65 ± 8.46	52.42 ± 8.34	52.87 ± 8.57
Sex			
• Female	48.52	50.02	47.08
• Male	51.48	49.98	52.92
Race			
• White	51.69	52.98	50.46
• Black	24.66	24.53	24.78
• Hispanic	12.01	10.26	13.69
• Other	11.64	12.23	11.07
RxRisk	9.33 ± 4.09	10.39 ± 3.39	8.32 ± 4.43
RxRisk (Weighted)	12.76 ± 7.03	13.82 ± 6.30	11.75 ± 7.53
CCI (Unweighted)	2.49 ± 1.95	2.47 ± 1.92	2.50 ± 1.97
CCI	3.54 ± 2.98	3.48 ± 2.89	3.60 ± 3.07

The results of the risk-adjusted analyses predicting beta-blocker use post MI discharge are presented in Table 56. All parameters included in the models, including age, sex, race/ethnicity, and RxRisk score were found to be significant predictors of the TZD adherence measure. The c-statistic was 0.646 and 0.645 for the classical logistic and hierarchical logistic regression models respectively, showing modest discriminative ability for both models. The odds ratio estimates for age, gender and RxRisk variables from the classical logistic regression model and the hierarchical logistic regression model with a random intercept were similar.

Female gender and Hispanic ethnicity were less likely to use beta-blockers post MI in both models. The effect of other two races on beta-blocker use was reversed in the two models with Blacks and other races being more (less) likely to be filling a beta-blocker within 30 days of discharge compared to whites in the classical (hierarchical) logistic regression models. An increase in RxRisk index by 1 was shown to increase the likelihood of beta-blocker use by approximately 15% in the two models. Age was a significant predictor, but with modest association with beta-blocker use.

Table 56: Odds ratio estimates of patient characteristics in the risk adjustment models

Baseline Characteristics	Classical Logistic Regression Model*	Hierarchical Logistic Regression Random Intercept Model*
	Point Estimate	Point Estimate
Age	0.985	0.986
Sex		
• Female vs. Male	0.853	0.848
Race		
• Black vs. White	1.014	0.984
• Hispanic vs. White	0.819	0.885
• Other vs. White	1.185	0.908
RxRisk	1.151	1.145

*significant at $p < 0.0001$

In the hierarchical logistic regression model with a random-intercept, the state-level variance component was estimated to be 0.1523 (SE: 0.04533). Testing the null hypothesis of no random effects using a likelihood ratio test based on residual pseudo-likelihood yielded a chi-square of 180.61 ($p<0.0001$) indicating the presence of random effect. Therefore, case-mix adjusted scores based on hierarchical logistic regression model with a random intercept are presented in Table 57. The residual intra class correlation coefficient (ρ) for the random intercept model was estimated to be 0.04424 which indicates that 4.42% of the unexplained variation after

controlling for patient level variables could be attributed to variation between states. State level explained more variation in MI1 measure compared to all other measures being studied.

Table 57 shows the agreement between the crude and case-mix adjusted scores. The case-mix adjusted MI1 measure scores ranged from 25.49% to 68.44%, a decrease in range compared to the crude estimates. Case-mix adjustment ranked the states differently with 14% (5) states ranked the same, 28% (10) states changing one position and 58% (21) changing more than two positions. There was good agreement in rankings based on the crude scores and case-mix adjusted scores based on the hierarchical logistic regression model with a random intercept (Kendall's τ_b=0.84).

Table 57: Agreement in ranks: Crude and case-mix adjusted MI1 measure scores[a]

State	Unadjusted Estimates		Case-Mix Adjusted Estimates		Score difference	Rank difference
	Score (%)	Rank	Score (%)	Rank		
Alabama	63.04	**5**	59.35	**7**	3.69	-2
Alaska	-	**-**	-	**-**	-	-
Arizona	58.72	**11**	60.72	**4**	-2.01	7
Arkansas	40.66	**29**	42.39	**29**	-1.73	0
California	47.57	**24**	50.55	**21**	-2.98	3
Colorado	62.96	**6**	64.15	**2**	-1.18	4
Connecticut	54.39	**16**	50.66	**20**	3.73	-4
Delaware	-	**-**	-	**-**	-	-
District of Columbia	49.09	**23**	52.86	**16**	-3.77	7
Florida	51.45	**19**	52.01	**17**	-0.56	2
Georgia	57.97	**12**	55.96	**12**	2.01	0
Hawaii	54.93	**15**	57.72	**10**	-2.79	5
Idaho	-	**-**	-	**-**	-	-
Illinois	50.63	**20**	49.49	**22**	1.14	-2
Indiana	38.79	**30**	36.52	**31**	2.27	-1
Iowa	26.83	**35**	25.49	**36**	1.34	-1
Kansas	45.45	**25**	42.63	**28**	2.82	-3
Kentucky	65.23	**3**	58.74	**8**	6.49	-5
Louisiana	57.72	**13**	54.82	**15**	2.90	-2

Table 57: Agreement in ranks: Crude and case-mix adjusted MI1 measure scores[a] (cont.)

State	Unadjusted Estimates		Case-Mix Adjusted Estimates		Score difference	Rank difference
	Score (%)	Rank	Score (%)	Rank		
Maine	-	-	-	-	-	-
Maryland	42.54	**28**	43.85	**27**	-1.32	1
Massachusetts	68.33	**1**	68.44	**1**	-0.10	0
Michigan	-	-	-	-	-	-
Minnesota	59.09	**9**	56.91	**11**	2.18	-2
Mississippi	37.96	**31**	41.23	**30**	-3.28	1
Missouri	50.21	**22**	48.15	**25**	2.06	-3
Montana	-	-	-	-	-	-
Nebraska	-	-	-	-	-	-
Nevada	32.35	**33**	33.92	**34**	-1.57	-1
New Hampshire	-	-	-	-	-	-
New Jersey	58.87	**10**	55.14	**14**	3.73	-4
New Mexico	60.61	**8**	58.33	**9**	2.28	-1
New York	32.32	**34**	34.61	**33**	-2.28	1
North Carolina	53.26	**17**	50.69	**19**	2.57	-2
North Dakota	-	-	-	-	-	-
Ohio	-	-	-	-	-	-
Oklahoma	44.97	**26**	44.49	**26**	0.48	0
Oregon	19.72	**36**	28.06	**35**	-8.34	1
Pennsylvania	42.92	**27**	48.69	**24**	-5.76	3
Rhode Island	-	-	-	-	-	-
South Carolina	50.55	**21**	51.37	**18**	-0.82	3
South Dakota	-	-	-	-	-	-
Tennessee	56.40	**14**	55.95	**13**	0.45	1
Texas	34.03	**32**	36.24	**32**	-2.20	0
Utah	-	-	-	-	-	-
Vermont	-	-	-	-	-	-
Virginia	52.63	**18**	49.03	**23**	3.60	-5
Washington	61.69	**7**	60.19	**5**	1.50	2
West Virginia	63.50	**4**	59.66	**6**	3.84	-2
Wisconsin	66.67	**2**	62.35	**3**	4.31	-1
Wyoming	-	-	-	-	-	-
[b]Kendall's τ_b	0.84					

[a]MI1 measure scores could be generated for 36 states. Rankings ranged from 1-36

[b]Kendall's τ_b is a nonparametric measure of association based on the number of concordances and discordances in rankings based on unadjusted and risk-adjusted scores

Additionally, states were classified into top (20%), medium and bottom (20%) performers based on the crude and case-mix adjusted scores. Results of the agreement in the grouping based on the unadjusted and risk-adjusted groupings are shown in Table 58a and Table 58b. Results based on the two risk adjustment models showed close to perfect agreement in classification of states into top, medium and bottom groups (κ=0.90). There was good agreement in classification of the states based on the crude and case-mix adjusted models when case-mix adjustment was conducted using both classical logistic regression (κ=0.80) and random intercept models (κ=0.90).

An alternate methodology was proposed for classification of states by identifying outliers as low (high) quality outliers if the score for a state was significantly lower (or higher) than the average score according to the 95% CI of the measure. The results of agreement in classification based on this methodology are included in Table A-15 (APPENDIX). Choropleth maps depicting the top, medium and bottom performing states based on the crude and case-mix adjustment scores estimated from the hierarchical logistic regression model are depicted in Figure 11. Additional choropleth maps illustrating the distribution of crude adherence scores are presented in Figure A-13 (APPENDIX).

Table 58a: Agreement in groups: Crude and case-mix adjusted MI1 measure scores

Groups Based on Crude Estimates	Groups Based on Case-mix Adjusted Estimates					
	Classical Logistic Regression Model			Hierarchical Logistic Regression Random Intercept Model		
	Bottom	Medium	Top	Bottom	Medium	Top
Bottom (~20%)	7	0	0	7	0	0
Medium (60%)	0	20	2	0	21	1
Top (~20%)	0	2	5	0	1	6
Percentage misclassified[a]	0%	9.1%	28.6%	0%	4.5%	14.2%
Cohen's κ[b]	0.80			0.90		

[a]Percentage misclassified was calculated for each risk adjustment method using the classification based on the risk adjustment method as the correct classification. [b]Cohen's Kappa coefficient is a measure of agreement between unadjusted performance ratings and ratings based on the risk adjustment method with higher values indicating greater agreement

Table 58b: Agreement in groups: Case-mix adjusted MI1 measure scores

Groups Based on Case-mix Adjusted Estimates	Hierarchical Logistic Regression Model		
Logistic Regression Model	Bottom	Medium	Top
Bottom	7	0	0
Medium	0	21	1
Top	0	1	6
Percentage misclassified[a]	0%	4.5%	14.2%
Cohen's κ[b]	0.90		

[a]Percentage misclassified was calculated for each risk adjustment method using the classification based on hierarchical logistic regression model as the correct classification. [b]Cohen's Kappa coefficient is a measure of agreement between unadjusted performance ratings and ratings based on the risk adjustment method with higher values indicating greater agreement.

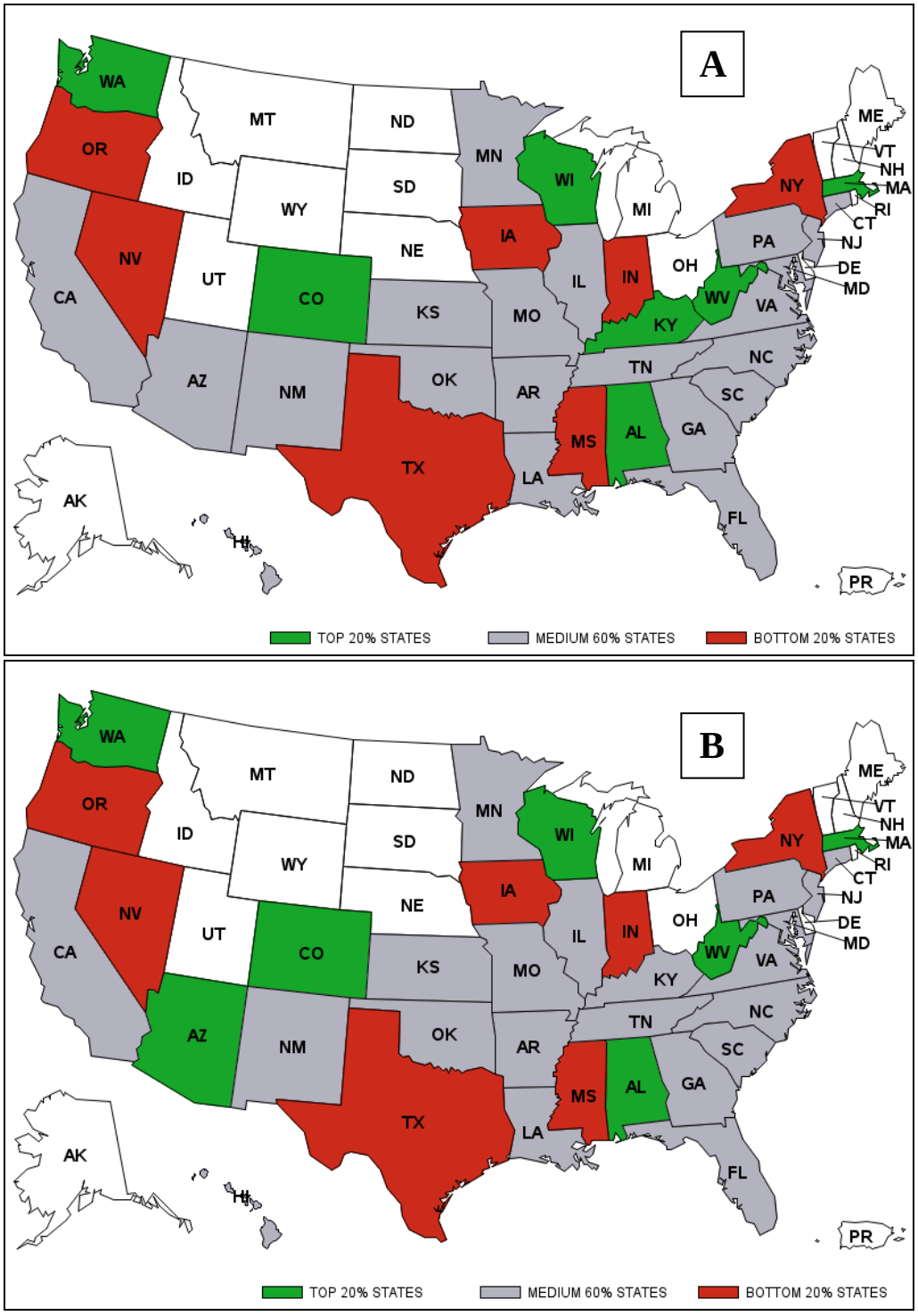

Figure 11. Interstate variations in MI1 measure:
(A) Unadjusted scores, (B) Case-mix adjusted scores

Persistence with beta-blockers in myocardial infarction patients (MI)

The MI measure assessing the proportion of patients persistent with beta-blockers after an acute myocardial infarction could be computed for 30 states. The number of patients eligible for the measure in each state and the proportion of patients persistent with beta-blocker therapy are shown in Table 59. The number of patients in the denominator ranged from 31 in Connecticut to 460 in California. The average MI score across all states was 58.9%, ranging from 40.5% of Medicaid patients being persistent with beta-blocker therapy in Arkansas to 71.1% in Indiana. States with scores above the national benchmark are highlighted in bold.

Table 59: Patients meeting the MI measure criteria by state

State	MI	
	#patients in the denominator	%patients in the numerator
Alabama	116	**63.79**
Alaska	-	-
Arizona	165	58.18
Arkansas	37	40.54
California	460	**65.87**
Colorado	34	50.00
Connecticut	31	**64.52**
Delaware	-	-
District of Columbia	-	-
Florida	213	**60.56**
Georgia	160	53.13
Hawaii	39	43.59
Idaho	-	-
Illinois	240	**62.08**
Indiana	83	**71.08**
Iowa	-	-
Kansas	-	-
Kentucky	212	58.02
Louisiana	142	52.11
Maine	-	-
Maryland	57	50.88
Massachusetts	123	56.91
Michigan	-	-
Minnesota	52	**63.46**

Table 59: Patients meeting the MI measure criteria by state (continued)

State	MI	
	#patients in the denominator	%patients in the numerator
Mississippi	52	50.00
Missouri	122	56.56
Montana	-	-
Nebraska	-	-
Nevada	-	-
New Hampshire	-	-
New Jersey	83	55.42
New Mexico	-	-
New York	267	**65.92**
North Carolina	229	53.71
North Dakota	-	-
Ohio	-	-
Oklahoma	67	50.75
Oregon	-	-
Pennsylvania	91	57.14
Rhode Island	-	-
South Carolina	92	41.30
South Dakota	-	-
Tennessee	163	53.99
Texas	227	56.39
Utah	-	-
Vermont	-	-
Virginia	70	58.57
Washington	95	**66.32**
West Virginia	127	**67.72**
Wisconsin	96	**64.58**
Wyoming	-	-
Total	3,945	58.94

Demographic characteristics and co-morbidity measures for the eligible patient population are described in Table 60. The mean age (± SD) of the study population was 52 ± 8 years and approximately 50% were females. The majority of the patients included in the MI measure was white (53%) followed by Black (24%), other races (12%) and Hispanic (10%). Non-persistent patients were slightly younger, and a greater proportion of non-persistent patients were male and Black compared to persistent patients. The mean comorbidity scores were slightly

higher in the persistent group compared to Medicaid beneficiaries' non-persistent with beta-blocker therapy post AMI.

Table 60: Characteristics of Medicaid patients eligible for MI measure

Patient Characteristics	Prevalence(%) or Mean ± SD		
	Overall	**Persistent**	**Non-persistent**
Age	52.41 ± 8.35	53.01 ± 7.99	51.56 ± 8.77
Sex			
• Female	49.94	51.91	47.10
• Male	50.06	48.09	52.90
Race			
• White	53.21	55.27	50.25
• Black	24.23	21.68	27.90
• Hispanic	10.34	10.49	10.12
• Other	12.22	12.56	11.73
RxRisk	10.38 ± 3.39	10.75 ± 3.33	9.84 ± 3.40
RxRisk (Weighted)	13.81 ± 6.29	14.25 ± 6.30	13.17 ± 6.23
CCI (Unweighted)	2.47 ± 1.92	2.57 ± 1.89	2.33 ± 1.95
CCI	3.47 ± 2.89	3.61 ± 2.84	3.28 ± 2.95

The results of the risk-adjusted analyses predicting persistence with beta-blocker therapy in post-MI Medicaid patients are presented in Table 61. Age and RxRisk were significant predictors of persistence with beta-blocker therapy in both models. The odds ratio estimates for these two variables from the classical logistic regression model and the hierarchical logistic regression model with a random intercept were found to be similar. Black race was a significant predictor of beta-blocker persistence in the classical logistic regression models, and it was marginally significant (p=0.054) in the hierarchical logistic regression model. Also, the c-statistic was 0.589 and 0.590 for the classical logistic and hierarchical logistic regression models respectively, showing weak discriminative ability for both models.

Table 61: Odds ratio estimates of patient characteristics in the risk adjustment models

Baseline Characteristics	Classical Logistic Regression Model	Hierarchical Logistic Regression Random Intercept Model
	Point Estimate	Point Estimate
Age	1.017*****	1.017*****
Sex		
• Female vs. Male	1.095	0.890
Race		
• Black vs. White	0.744*****	0.792**
• Hispanic vs. White	0.997	0.929
• Other vs. White	0.972	1.054
RxRisk	1.073*****	1.075*****

*significant at $p < 0.001$; **marginally significant $p=0.054$

In the hierarchical logistic regression model with a random-intercept, the state-level variance component was estimated to be 0.03971 (SE: 0.01987). Testing the null hypothesis of no random effects using a likelihood ratio test based on residual pseudo-likelihood yielded a chi-square of 15.14 ($p<0.0001$) indicating the presence of random effect. Case-mix adjusted scores based on hierarchical logistic regression model with a random intercept are presented in Table 62. The residual intra class correlation coefficient (ρ) for the random intercept model was estimated to be 0.01193 which indicates that 1.19% of the unexplained variation after controlling for patient level variables could be attributed to variation between states.

Table 62 shows the agreement between the crude and case-mix adjusted scores. The case-mix adjusted MI measure scores ranged from 42.05% to 66.38%, a decrease in range compared to the crude estimates. Case-mix adjustment ranked the states differently with 33% (10) states changing one position and 57% (17) changing more than two positions. There was good agreement in rankings based on the crude scores and case-mix adjusted scores based on the hierarchical logistic regression model with a random intercept (Kendall's $\tau_b=0.82$).

Table 62: Agreement in ranks: Crude and case-mix adjusted MI measure scores[a]

State	Unadjusted Estimates		Case-Mix Adjusted Estimates		Score difference	Rank difference
	Score (%)	Rank	Score (%)	Rank		
Alabama	63.79	**8**	62.30	**9**	1.49	-1
Alaska	-	**-**	-	**-**	-	-
Arizona	58.18	**13**	57.90	**12**	0.28	1
Arkansas	40.54	**30**	43.18	**28**	-2.64	2
California	65.87	**5**	66.31	**2**	-0.44	3
Colorado	50.00	**26**	48.48	**27**	1.52	-1
Connecticut	64.52	**7**	60.97	**10**	3.54	-3
Delaware	-	**-**	-	**-**	-	-
District of Columbia	-	**-**	-	**-**	-	-
Florida	60.56	**11**	59.54	**11**	1.02	0
Georgia	53.13	**22**	51.78	**24**	1.34	-2
Hawaii	43.59	**28**	42.76	**29**	0.83	-1
Idaho	-	**-**	-	**-**	-	-
Illinois	62.08	**10**	62.71	**8**	-0.62	2
Indiana	71.08	**1**	66.38	**1**	4.70	0
Iowa	-	**-**	-	**-**	-	-
Kansas	-	**-**	-	**-**	-	-
Kentucky	58.02	**14**	53.86	**19**	4.16	-5
Louisiana	52.11	**23**	50.76	**25**	1.35	-2
Maine	-	**-**	-	**-**	-	-
Maryland	50.88	**24**	53.00	**23**	-2.12	1
Massachusetts	56.91	**16**	56.73	**14**	0.18	2
Michigan	-	**-**	-	**-**	-	-
Minnesota	63.46	**9**	62.89	**7**	0.57	2
Mississippi	50.00	**26**	53.92	**18**	-3.92	8
Missouri	56.56	**17**	54.62	**16**	1.93	1
Montana	-	**-**	-	**-**	-	-
Nebraska	-	**-**	-	**-**	-	-
Nevada	-	**-**	-	**-**	-	-
New Hampshire	-	**-**	-	**-**	-	-
New Jersey	55.42	**19**	53.42	**21**	2.00	-2
New Mexico	-	**-**	-	**-**	-	-
New York	65.92	**4**	65.75	**3**	0.17	1
North Carolina	53.71	**21**	53.21	**22**	0.50	-1
North Dakota	-	**-**	-	**-**	-	-

Table 62: Agreement in ranks: Crude and case-mix adjusted MI measure scores[a] (continued)

State	Unadjusted Estimates		Case-Mix Adjusted Estimates		Score difference	Rank difference
	Score (%)	Rank	Score (%)	Rank		
Ohio	-	-	-	-	-	-
Oklahoma	50.75	**25**	49.93	**26**	0.81	-1
Oregon	-	-	-	-	-	-
Pennsylvania	57.14	**15**	54.44	**17**	2.70	-2
Rhode Island	-	-	-	-	-	-
South Carolina	41.30	**29**	42.05	**30**	-0.75	-1
South Dakota	-	-	-	-	-	-
Tennessee	53.99	**20**	53.61	**20**	0.38	0
Texas	56.39	**18**	57.37	**13**	-0.98	5
Utah	-	-	-	-	-	-
Vermont	-	-	-	-	-	-
Virginia	58.57	**12**	56.00	**15**	2.57	-3
Washington	66.32	**3**	63.81	**5**	2.50	-2
West Virginia	67.72	**2**	63.01	**6**	4.71	-4
Wisconsin	64.58	**6**	64.52	**4**	0.07	2
Wyoming	-	-	-	-	-	-
[b]Kendall's τ_b	0.82					

[a]MI measure scores could be generated for 30 states. Rankings ranged from 1-30

[b]Kendall's τ_b is a nonparametric measure of association based on the number of concordances and discordances in rankings based on unadjusted and risk-adjusted scores

Additionally, states were classified into top (20%), medium and bottom (20%) performers based on the crude and case-mix adjusted scores. Results of the agreement in the grouping based on the unadjusted and risk-adjusted groupings are shown in Table 63a and Table 63b. Results based on the two risk adjustment models showed perfect agreement in classification of states into top, medium and bottom groups (κ=1.00). There was good agreement in classification of the states based on the crude and case-mix adjusted models when case-mix adjustment was conducted using both classical logistic regression (κ=0.88) and random intercept models (κ=0.88).

An alternate methodology was proposed for classification of states by identifying outliers as low (high) quality outliers if the score for a state was significantly lower (or higher) than the average score according to the 95% CI of the measure. The results of agreement in classification based on this methodology are included in Table A-16 (APPENDIX). Choropleth maps depicting the top, medium and bottom performing states based on the crude and case-mix adjustment scores estimated from the hierarchical logistic regression model are depicted in Figure 12. Additional choropleth maps illustrating the distribution of crude adherence scores are presented in Figure A-14 (APPENDIX).

Table 63a: Agreement in groups: Crude and case-mix adjusted MI measure scores

Groups Based on Crude Estimates	Groups Based on Case-mix Adjusted Estimates					
	Classical Logistic Regression Model			Hierarchical Logistic Regression Random Intercept Model		
	Bottom	Medium	Top	Bottom	Medium	Top
Bottom (~20%)	5	1	0	5	1	0
Medium (60%)	1	17	0	1	17	0
Top (~20%)	0	0	6	0	0	6
Percentage misclassified[a]	16.7%	5.5%	0%	16.7%	5.5%	0%
Cohen's κ[b]	0.88			0.88		

[a]Percentage misclassified was calculated for each risk adjustment method using the classification based on the risk adjustment method as the correct classification. [b]Cohen's Kappa coefficient is a measure of agreement between unadjusted performance ratings and ratings based on the risk adjustment method with higher values indicating greater agreement

Table 63b: Agreement in groups: Case-mix adjusted MI measure scores

Groups Based on Case-mix Adjusted Estimates	Hierarchical Logistic Regression Model		
Logistic Regression Model	Bottom	Medium	Top
Bottom	6	0	0
Medium	0	18	0
Top	0	0	6
Percentage misclassified[a]	0%	0%	0%
Cohen's κ[b]	1.00		

[a]Percentage misclassified was calculated for each risk adjustment method using the classification based on hierarchical logistic regression model as the correct classification. [b]Cohen's Kappa coefficient is a measure of agreement between unadjusted performance ratings and ratings based on the risk adjustment method with higher values indicating greater agreement.

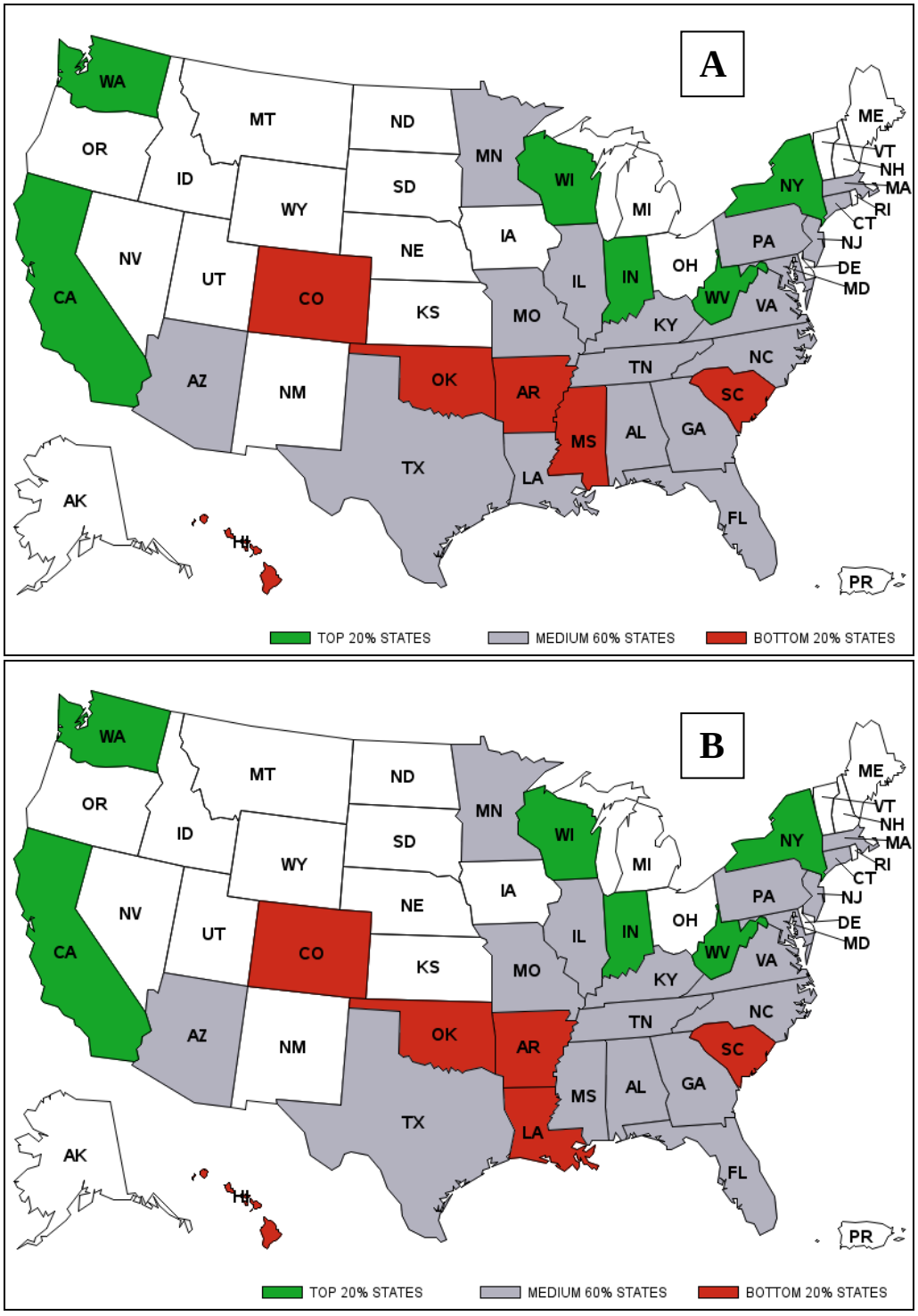

Figure 12. Interstate variations in MI measure:
(A) Unadjusted scores, (B) Case-mix adjusted scores

Composite Measure

A composite measure of medication use-related quality was computed for 42 states and the District of Columbia. Demographic characteristics and co-morbidity measures for the 1,536,753 eligible patients are described in Table 64. The majority of the study population was comprised of women (68%) and the mean age (± SD) was 46.18 ± 11.94 years. Patients who did not receive good quality of care on any of the medication use-related quality measures were younger, had less comorbidity burden and a greater proportion were female, and of Black, Hispanic race/ethnicity compared to those receiving good quality care.

Table 64: Characteristics of Medicaid patients eligible for the composite measure

Patient Characteristics	Prevalence(%) or Mean ± SD		
	Overall	**Good Quality***	**Poor Quality****
Age	46.18 ± 11.94	47.78 ± 11.62	43.92 ± 12.01
Sex			
• Female	68.01	66.65	69.93
• Male	31.99	33.35	30.07
Race			
• White	45.06	47.84	41.12
• Black	25.03	22.84	28.15
• Hispanic	16.77	15.60	18.42
• Other	13.14	13.72	12.31
RxRisk	5.95 ± 3.00	6.46 ± 3.10	5.23 ± 2.71
RxRisk (Weighted)	7.87 ± 5.12	8.50 ± 5.24	6.99 ± 4.82
CCI (Unweighted)	1.0 ± 1.14	1.13 ± 1.18	0.81 ± 1.05
CCI	1.56 ± 1.87	1.75 ± 1.91	1.30 ± 1.77

*Good quality of care is patients qualifying for at least one of the 11 study measures (meeting the adherence threshold for the PDC measures, being persistent for the MI, AD measures and filling a recommended prescription for the CAD and ICS measures). **Poor Quality is patients not qualifying for the numerator for any of the 11 measures included in the study.

The results of the risk-adjusted analyses predicting good quality of care accounting for patient case-mix and opportunity mix are presented in Table 65. Two hierarchical logistic regression models were explored, with and without accounting for nesting of observations within patients. In the two-level model not accounting for observations nested within patients, age, sex, race/ethnicity and RxRisk score were found to be significant predictors of medication utilization.

To account for observations nested within patients and patients nested within states, a three-level hierarchical logistic regression model with patient and state levels modeled as random intercepts (G-side random effects) was explored. However, the model did not converge. Therefore, the patient level was modeled as a residual random effect (patient as R-side random effects) and the state level as a G-side random effect. All case-mix and opportunity mix variables were significant predictors of medication utilization behavior. Race/ethnicity was a strong predictor of medication use in both models, with Blacks and Females being associated with significantly lesser odds of receiving good quality of care compared to Whites and males, respectively. Hispanics were nearly 35% less likely to receive good quality of care compared to Whites in both models. The c-statistic was 0.637 and 0.635 for the two-level and three-level hierarchical logistic regression models respectively, showing modest discriminative ability for both models.

Table 65: Odds ratio estimates of patient characteristics in the risk adjustment models

Baseline Characteristics	Hierarchical Logistic Regression Models	
	Two-level	Three-level
	Point Estimate	Point Estimate*
Age	1.025*	1.023
Sex		
• Female vs. Male	0.861*	0.866
Race		
• White vs. Other	1.084*	1.073
• Black vs. Other	0.704*	0.706
• Hispanic vs. Other	0.743*	0.748
RxRisk	1.049*	1.051
Measure		
• ACEI/ARB vs. MI	1.023	0.886
• BB vs. MI	0.888*	0.763
• CCB vs. MI	0.972	0.838
• BIGU vs. MI	0.751*	0.629
• SU vs. MI	0.799*	0.667
• TZD vs. MI	0.632*	0.519
• STAT vs. MI	0.794*	0.632
• AD vs. MI	0.535*	0.504
• CAD vs. MI	1.322*	1.411
• ICS vs. MI	2.879*	2.479

*significant at $p < 0.001$

In the two-level hierarchical logistic regression model, the state-level variance component was estimated to be 0.02685 (SE: 0.005910). Testing the null hypothesis of no random effects using a likelihood ratio test based on residual pseudo-likelihood yielded a chi-square of 11689.9 ($p<0.0001$) indicating the presence of random effect. The residual intra class correlation coefficient (ρ) for the random intercept model was estimated to be 0.00810 which indicates that 0.81% of the unexplained variation after controlling for patient case-mix variables and opportunity mix could be attributed to variation between states.

In the three-level hierarchical logistic regression model, the state-level variance component was estimated to be 0.02735 (SE: 0.006049) and the patient-level variance component was estimated to be 0.3972 (SE: 0.000903). Testing the null hypothesis of no random effects using a likelihood ratio test based on residual pseudo-likelihood yielded a chi-square of 381589 ($p<0.0001$) indicating the presence of random effects. When correlation of observations within patients was accounted for, the residual intra class correlation coefficient (ρ) was estimated to be 0.00824 which indicates that 0.82% of the unexplained variation in the composite score could be attributed to variation between states. A comparison of ranks and groups based on the case-mix adjusted scores using both the two-level and three-level models are presented in Tables 66, 67a and 67b.

Table 66 shows the agreement between the crude and case-mix adjusted scores (both two-level and three-level models). The case-mix adjusted composite measure scores from the three-level model ranged from 45.99% to 63.24%, a slight decrease in range compared to the crude estimates. Case-mix adjustment ranked the states differently with 81% (35) states changing more than two positions. There was good agreement in rankings based on the crude scores and case-mix adjusted scores based on the two-level (Kendall's τ_b=0.75) and three-level (Kendall's τ_b=0.74) hierarchical logistic regression models.

Table 66: Agreement in ranks: Crude and case-mix adjusted composite measure scores[a]

State	Unadjusted Estimates		Case-Mix Adjusted Estimates[b]			Score difference	Rank difference
	Score (%)	Rank	Score (%)	Rank 3 Level	Rank 2 Level		
Alabama	51.40	**23**	52.80	**23**	**23**	-1.40	0
Alaska	-	**-**	-	**-**	**-**	-	-
Arizona	45.32	**40**	48.00	**39**	**39**	-2.68	1
Arkansas	46.65	**36**	48.99	**36**	**37**	-2.33	0
California	46.16	**38**	49.61	**34**	**34**	-3.45	4
Colorado	45.93	**39**	47.18	**41**	**41**	-1.26	-2
Connecticut	58.81	**5**	60.03	**2**	**2**	-1.23	3
Delaware	50.78	**26**	54.16	**19**	**20**	-3.37	7
District of Columbia	47.85	**34**	52.51	**25**	**25**	-4.66	9
Florida	53.37	**15**	53.62	**22**	**22**	-0.25	-7
Georgia	45.13	**41**	46.54	**42**	**42**	-1.42	-1
Hawaii	50.61	**27**	49.31	**35**	**35**	1.29	-8
Idaho	60.61	**2**	59.56	**3**	**3**	1.05	-1
Illinois	52.34	**18**	55.75	**9**	**9**	-3.41	9
Indiana	49.13	**32**	48.49	**38**	**38**	0.64	-6
Iowa	58.41	**7**	58.04	**5**	**5**	0.37	2
Kansas	55.87	**12**	55.38	**10**	**10**	0.49	2
Kentucky	57.69	**8**	54.28	**18**	**18**	3.41	-10
Louisiana	48.67	**33**	51.61	**28**	**29**	-2.94	5
Maine	-	**-**	-	**-**	**-**	-	-
Maryland	46.99	**35**	50.41	**33**	**33**	-3.42	2
Massachusetts	51.53	**22**	50.87	**31**	**32**	0.66	-9
Michigan	-	**-**	-	**-**	**-**	-	-
Minnesota	50.01	**30**	52.16	**27**	**26**	-2.15	3
Mississippi	40.58	**43**	45.99	**43**	**43**	-5.41	0
Missouri	55.47	**14**	54.93	**11**	**11**	0.55	3
Montana	-	**-**	-	**-**	**-**	-	-
Nebraska	57.59	**9**	58.46	**4**	**4**	-0.87	5
Nevada	56.12	**10**	54.79	**12**	**12**	1.33	-2
New Hampshire	59.74	**3**	57.32	**8**	**8**	2.42	-5
New Jersey	50.82	**25**	52.52	**24**	**24**	-1.71	1
New Mexico	50.05	**29**	52.31	**26**	**27**	-2.26	3
New York	51.33	**24**	54.30	**17**	**21**	-2.96	7
North Carolina	51.83	**19**	54.14	**20**	**19**	-2.32	-1

Table 66: Agreement in ranks: Crude and case-mix adjusted composite measure scores[a] (continued)

State	Unadjusted Estimates		Case-Mix Adjusted Estimates[b]			Score difference	Rank difference
	Score (%)	Rank	Score (%)	Rank 3 Level	Rank 2 Level		
North Dakota	-	-	-	-	-	-	-
Ohio	-	-	-	-	-	-	-
Oklahoma	49.49	**31**	48.93	**37**	**36**	0.56	-6
Oregon	52.45	**17**	51.35	**29**	**28**	1.10	-12
Pennsylvania	59.66	**4**	57.81	**6**	**6**	1.85	-2
Rhode Island	51.82	**20**	54.60	**14**	**14**	-2.77	6
South Carolina	46.58	**37**	50.79	**32**	**31**	-4.21	5
South Dakota	-	-	-	-	-	-	-
Tennessee	45.02	**42**	47.19	**40**	**40**	-2.17	2
Texas	51.66	**21**	54.67	**13**	**13**	-3.01	8
Utah	50.36	**28**	51.18	**30**	**30**	-0.83	-2
Vermont	64.55	**1**	63.24	**1**	**1**	1.32	0
Virginia	55.77	**13**	57.43	**7**	**7**	-1.66	6
Washington	55.93	**11**	54.57	**15**	**15**	1.36	-4
West Virginia	58.73	**6**	54.46	**16**	**16**	4.27	-10
Wisconsin	52.46	**16**	54.09	**21**	**17**	-1.63	-5
Wyoming	-	-	-	-	-	-	-
[c]Kendall's τ_b				0.74	0.75		

[a]Composite measure measure scores were generated for 43 states. Rankings ranged from 1-43
[b]scores are from the three-level hierarchical logistic regression model with state (G-side) and patient (R-side) random effects
[c]Kendall's τ_b is a nonparametric measure of association based on the number of concordances and discordances in rankings based on unadjusted and risk-adjusted scores

Additionally, states were classified into top (20%), medium and bottom (20%) performers based on the crude and case-mix adjusted scores. Results of the agreement in the grouping based on the unadjusted and risk-adjusted groupings are shown in Table 67a and Table 67b. Results based on the two risk adjustment models showed perfect agreement in classification of states into top, medium and bottom groups (κ=1.00). There was modest agreement in classification of the states based on the crude and case-mix adjusted models when case-mix

adjustment was conducted using two-level (κ=0.65) and three-level (κ=0.65) hierarchical logistic regression models.

An alternate methodology was proposed for classification of states by identifying outliers as low (high) quality outliers if the score for a state was significantly lower (or higher) than the average score according to the 95% CI of the measure. However, the distribution of scores for the 43 states is leptokurtic i.e., concentrated about the mean, therefore not conducive to identifying outliers. The results of agreement in classification based on this methodology are included in Table A-17 (APPENDIX). Choropleth maps depicting the top, medium and bottom performing states based on the crude and case-mix adjustment scores estimated from the three-level hierarchical logistic regression model are depicted in Figure 13.

Table 67a: Agreement in groups: Crude and case-mix adjusted composite measure scores

Groups Based on Crude Estimates	Groups Based on Case-mix Adjusted Estimates					
	Hierarchical Logistic Regression Two-Level Model			Hierarchical Logistic Regression Three-Level Model		
	Bottom	Medium	Top	Bottom	Medium	Top
Bottom (~20%)	6	2	0	6	2	0
Medium (60%)	2	23	2	2	23	2
Top (~20%)	0	2	6	0	2	6
Percentage misclassified[a]	25%	14.8%	25%	25%	14.8%	25%
Cohen's κ[b]	0.65			0.65		

[a]Percentage misclassified was calculated for each risk adjustment method using the classification based on the risk adjustment method as the correct classification. [b]Cohen's Kappa coefficient is a measure of agreement between unadjusted performance ratings and ratings based on the risk adjustment method with higher values indicating greater agreement

Table 67b: Agreement in groups: Case-mix adjusted composite measure scores

Groups Based on Case-mix Adjusted Estimates	Three-Level Model		
Two-Level Model	Bottom	Medium	Top
Bottom	8	0	0
Medium	0	27	0
Top	0	0	8
Percentage misclassified[a]	0%	0%	0%
Cohen's κ[b]	1.00		

[a]Percentage misclassified was calculated for each risk adjustment method using the classification based on the three-level risk adjustment method as the correct classification. [b]Cohen's Kappa coefficient is a measure of agreement between unadjusted performance ratings and ratings based on the risk adjustment method with higher values indicating greater agreement

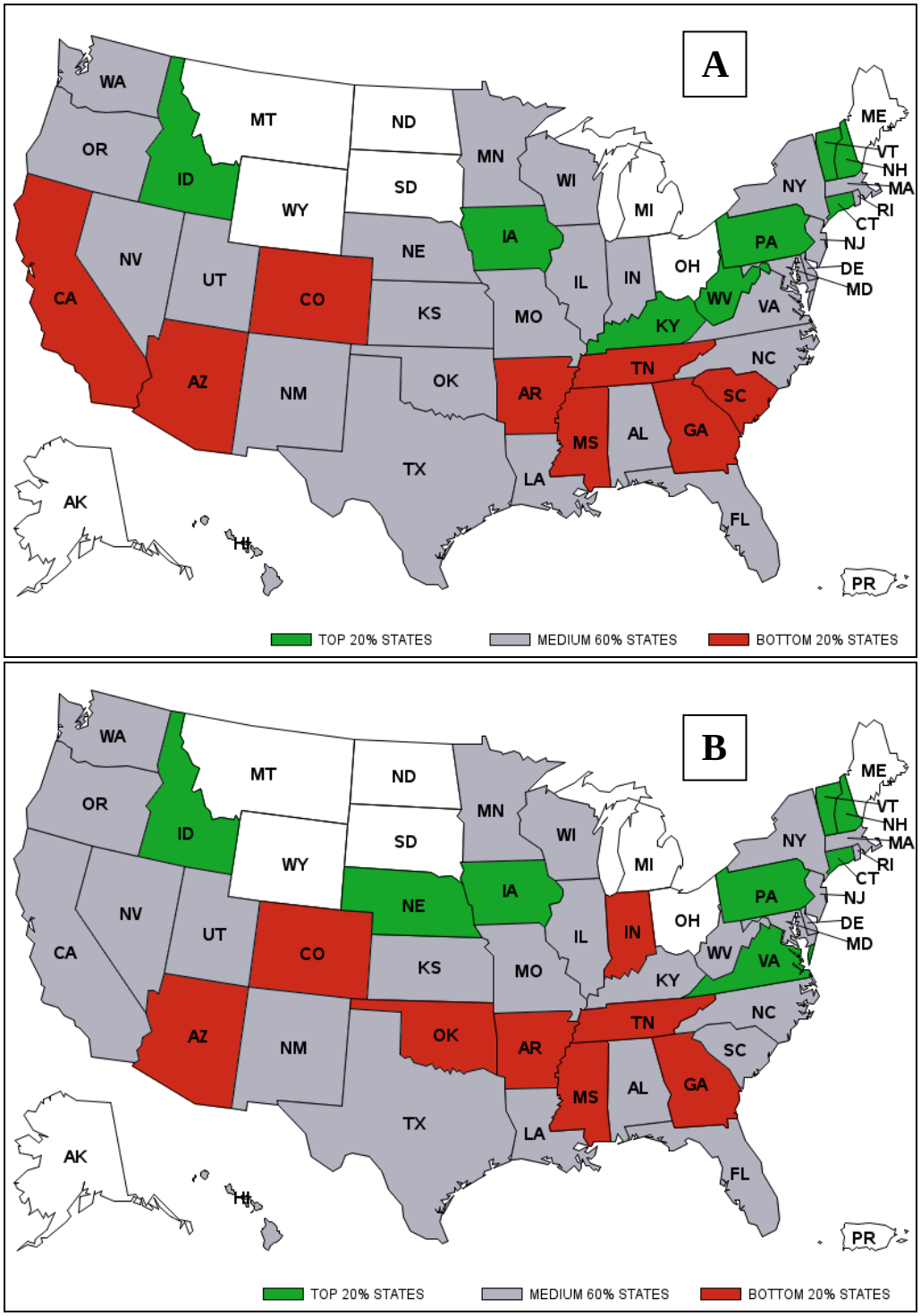

Figure 13. Interstate variations in composite measure: (A) Unadjusted scores, (B) Case-mix adjusted scores

A summary of the performance measures included in the study is presented in Table 68. Univariate statistics and coefficient of variation describing the variation in the 13 measures and the composite measure are tabulated. National benchmarks for Medicaid use-related measures for the year 2007 ranged from 31.5% for the chronic phase persistence with antidepressants measure to 66.8% for the ICS measure. There was substantial variation in the 13 measures being studied with coefficient of variation ranging from 6.7 for the ICS measure to 20.5 for the MI1 measure. The best performing state Medicaid programs also had significant room for improvement across all measures.

Table 68: Distribution of state performance scores by measure type

Measure	**2007** (%)							
	Mean	Min	10th percentile	Median	90th percentile	Max	SD	CV
PDC (Proportion of patients 18 years and older who met PDC threshold of 80 percent per county)								
ACEI/ARB	54.41	45.53	48.12	54.88	59.78	65.24	4.52	8.31
BB	50.69	35.82	44.85	50.99	56.81	60.62	5.16	10.18
CCB	53.28	40.46	48.59	52.81	58.05	65.43	4.45	8.36
Biguanides	45.72	35.77	40.12	46.00	50.63	54.37	4.20	9.19
Sulfonylureas	48.94	38.68	43.83	49.51	55.34	57.36	4.34	8.88
TZDs	44.54	33.22	39.04	44.52	49.99	54.61	4.65	10.43
Statins	51.50	42.10	45.24	50.99	58.88	64.05	5.15	10.00
AD Acute (Proportion of Medicaid patients 18-65 years newly diagnosed with MDD and prescribed antidepressant therapy persistent with the therapy for at least 3 months)								
AD Acute	52.80	40.08	48.92	52.41	57.26	64.44	4.16	7.89
AD Chronic (Proportion of Medicaid patients 18-65 years newly diagnosed with MDD and prescribed antidepressant therapy persistent with the therapy for at least 6 months)								
AD Chronic	31.53	13.41	26.53	31.86	36.40	42.49	4.90	15.54
CAD (Percent of patients 18-65 years of age diagnosed with CAD that received at least one prescription for a statin medication)								
CAD	63.04	35.47	58.70	63.22	68.89	73.49	5.85	9.29
ICS (Percent of asthma patients 18-50 years of age with persistent asthma that received at least one prescription for an inhaled corticosteroid or similar medication)								
ICS	66.81	57.55	60.60	67.15	73.06	75.57	4.46	6.67
MI (Proportion of Medicaid patients 18-65 years with MI that filled a beta-blocker prescription within 30 days of discharge persistent with therapy for at least 6 months)								
MI	56.17	42.05	45.83	55.31	65.13	66.38	6.90	12.29
MI1 (Proportion of Medicaid patients 18-65 years with MI that filled a beta-blocker prescription within 30 days of discharge)								
MI1	50.04	25.49	34.61	51.03	60.72	68.44	10.27	20.53
Composite	53.12	45.99	48.00	53.62	58.04	63.24	3.90	7.34

To assess if performance on one measure was correlated with performance on another, states were ranked on all measures and the correlations between the 11 medication use-related measures and the composite measure were assessed using Kendall's τ_b. Kendall's τ_b is a nonparametric rank correlation measure based on the number of concordances and discordances in rankings.

A summary of state rankings based on the various performance measures and the correlation between rankings on these measures is provided in Table 69. There was significant correlation between ranks based on the seven PDC measures included in the study, with τ_b values ranging from 0.54 to 0.73. However, there was no significant correlation in ranking of states based on the adherence/persistence measures (PDC, AD) and the standard of care measures (CAD, ICS). All measure based ranks except the MI measure based ranks were correlated with the composite measure. However, the τ_b values indicate that there is good agreement only between the composite measure based ranks and the PDC measure based ranks.

Table 69: Summary of agreement between state rankings on medication use-related performance measures

State	Rankings by Measure											
	ACEI/ ARB	BB	CCB	BIGU	SU	TZD	STAT	AD	CAD	ICS	MI	Composite
Alabama	25	28	18	29	16	33	23	20	33	15	9	23
Alaska	-	-	-	-	-	-	-	-	-	-	-	-
Arizona	39	38	39	35	38	37	40	7	25	36	12	39
Arkansas	21	43	27	33	33	13	34	36	35	33	28	36
California	37	35	38	37	40	31	37	22	15	10	2	34
Colorado	43	37	41	41	41	40	39	1	23	41	27	41
Connecticut	4	3	4	7	4	6	3	14	8	4	10	2
Delaware	19	21	35	24	12	23	30	15	5	1	-	19
District of Columbia	28	32	21	11	19	42	27	35	29	22	-	25
Florida	22	17	28	12	20	24	21	28	41	13	11	22
Georgia	42	40	42	39	39	35	41	41	32	35	24	42
Hawaii	41	34	36	27	32	18	15	43	40	32	29	35
Idaho	2	2	2	3	2	2	2	19	24	37	-	3
Illinois	16	11	17	13	7	20	19	32	3	9	8	9
Indiana	38	33	40	40	36	39	33	29	28	42	1	38
Iowa	7	4	10	6	14	3	4	3	12	23	-	5
Kansas	11	9	9	18	29	8	8	21	34	31	-	10
Kentucky	12	15	13	19	13	12	17	37	21	38	19	18
Louisiana	29	30	24	34	28	28	32	38	19	16	25	28
Maine	-	-	-	-	-	-	-	-	-	-	-	-
Maryland	31	29	43	20	21	38	28	27	37	28	23	33
Massachusetts	32	26	26	32	31	34	38	33	6	34	14	31
Michigan	-	-	-	-	-	-	-	-	-	-	-	-
Minnesota	33	22	37	31	34	10	24	24	9	21	7	27

Table 69: Summary of agreement between state rankings on medication use-related performance measures (continued)

State	Rankings by Measure											
	ACEI/ ARB	BB	CCB	BIGU	SU	TZD	STAT	AD	CAD	ICS	MI	Composite
Mississippi	40	42	34	42	43	43	42	42	31	7	18	43
Missouri	13	19	25	17	8	14	9	26	11	25	16	11
Montana	-	-	-	-	-	-	-	-	-	-	-	-
Nebraska	9	5	7	16	9	5	5	5	17	3	-	4
Nevada	15	12	12	15	3	11	16	13	42	30	-	12
New Hampshire	8	8	3	2	11	9	6	6	16	17	-	8
New Jersey	26	24	29	22	24	30	35	31	26	5	21	24
New Mexico	27	27	20	21	26	36	25	16	27	39	-	26
New York	23	23	22	9	10	32	29	9	7	6	3	17
North Carolina	24	18	11	26	22	16	18	23	18	14	22	20
North Dakota	-	-	-	-	-	-	-	-	-	-	-	-
Ohio	-	-	-	-	-	-	-	-	-	-	-	-
Oklahoma	36	41	23	43	42	26	26	17	39	29	26	37
Oregon	10	14	15	5	15	21	10	12	43	27	-	29
Pennsylvania	3	6	8	10	6	4	11	8	20	12	17	6
Rhode Island	6	20	6	28	25	19	36	25	2	11	-	14
South Carolina	30	39	31	30	30	25	31	39	10	24	30	32
South Dakota	-	-	-	-	-	-	-	-	-	-	-	-
Tennessee	35	36	33	36	35	41	43	40	30	43	20	40
Texas	18	25	14	23	27	15	12	4	36	2	13	13
Utah	34	31	32	38	37	29	22	18	14	20	-	30
Vermont	1	1	1	1	1	1	1	2	1	8	-	1
Virginia	5	7	5	4	5	7	7	30	38	18	15	7
Washington	14	16	16	8	18	22	13	10	13	40	5	15

Table 69: Summary of agreement between state rankings on medication use-related performance measures (continued)

State	Rankings by Measure											
	ACEI/ ARB	**BB**	**CCB**	**BIGU**	**SU**	**TZD**	**STAT**	**AD**	**CAD**	**ICS**	**MI**	**Compo site**
West Virginia	17	10	19	25	23	17	20	34	22	19	6	16
Wisconsin	20	13	30	14	17	27	14	11	4	26	4	21
Wyoming	-	-	-	-	-	-	-	-	-	-	-	-
	[a]Kendall's τ_b											
ACEI/ARB	1.00	0.73*	0.73*	0.64*	0.69*	0.65*	0.64*	0.31**	0.17	0.23**	0.17	0.80*
BB		1.00	0.59*	0.65*	0.67*	0.62*	0.67*	0.33**	0.21**	0.21**	0.33**	0.78*
CCB			1.00	0.54*	0.55*	0.56*	0.60*	0.28**	0.11	0.20	0.08	0.67*
BIGU				1.00	0.73*	0.46*	0.61*	0.30**	0.10	0.18	0.23	0.64*
SU					1.00	0.46*	0.56*	0.25**	0.17	0.23**	0.20	0.69*
TZD						1.00	0.66*	0.24**	0.13	0.19	0.01	0.65*
STAT							1.00	0.34**	0.05	0.14	0.12	0.68*
AD								1.00	0.15	0.13	0.27**	0.35**
CAD									1.00	0.19	0.28**	0.23**
ICS										1.00	0.14	0.33**
MI											1.00	0.24
Composite												1.00

[a]Kendall's τ_b is a nonparametric measure of association based on no. of concordances and discordances in rankings based on unadjusted and risk-adjusted scores
*Significant at P<0.0001level
**Significant at P<0.05 level

A very small proportion of variation in the study measures was explained by the state level random effect. A summary of the state-level contribution to the various measures being studied is shown in Table 70. The intraclass correlation coefficients from the unconditional means models (including only state as a random effect) were very small for all measures, indicating that there was only a moderate degree of clustering within states, and even after controlling for state random effects; considerable unexplained variability in the measures exists.

Table 70: Summary of state-level effects on performance measures

Measure	ICC[a]	Residual ICC[b]
ACEI/ARB	0.0190	0.0110
BB	0.0209	0.0136
CCB	0.0193	0.0097
BIGU	0.0153	0.0090
SU	0.0153	0.0087
TZD	0.0178	0.0098
STAT	0.0228	0.0137
AD Acute	0.0076	0.0068
AD Chronic	0.0142	0.0147
CAD	0.0410	0.0182
ICS	0.0122	0.0122
MI	0.0127	0.0119
MI1	0.0571	0.0442
Composite	0.0129	0.0081

[a]intraclass correlation coefficients from the unconditional means models (including only state as a random effect)
[b]residual intraclass correlation coefficients after controlling for patient-level variables

The amount of variation in the performance measures explained by the state effect was highest for the MI1 measure, with state level explaining 5.7% of variation in the measure. Amount of variation in CAD measure attributable to the state level was also substantial at 4.1%. State level contributed to explaining approximately 1.5 – 2% of the PDC measures. Models adjusting for patients' level factors including age, gender, race/ethnicity, and comorbidity burden decreased the amount of variation explained by the state random effect only slightly, as represented by the residual intraclass correlation coefficients shown in Table 70.

CHAPTER V – DISCUSSION AND CONCLUSIONS

The U.S. health care system is moving towards a value-driven payment model. Value entails provision of high quality care at low costs. Historically, Medicaid and Medicare programs relied on fee-for-service payment models that did not factor in the quality of care provided. Today, several government agencies, private health plans and large employer groups are relying on pay-for-performance models that incentivize providers based on the quality of care provided. Some of the noted federal pay-for-performance initiatives include the bonus payments provided to the top performing hospitals and nursing homes by CMS as a part of the Medicare value based purchasing demonstration projects.[113,114] Similar strategies can be employed by Medicaid to improve quality of care delivered to its enrollees.

Public reporting of performance measures is another important method adopted by the Federal government to reduce variation in quality and improve overall quality of care delivered. Over the past decade, reports of hospital, nursing home, and health plan quality have been made publicly available for consumers of healthcare to make informed decisions.[1,2,3] CMS is currently using public reporting in conjunction with pay-for-performance to improve quality of care provided by hospitals, nursing homes and Medicare Advantage plans. A few of the measures used by CMS to assess hospital quality and a considerable number of measures in the HEDIS measure set from NCQA used to determine health plan and managed care organization quality

are related to medication use.[1,14] In addition, PQA has developed a comprehensive list of medication use-related measures that can be used to assess quality of pharmacies and health care plans.[10,11] CMS currently uses some of these medication use-related quality measures (e.g., adherence with medications and high-risk medication use in elderly) to incentivize top performing Medicare Advantage plans providing drug benefits via a star rating system. These measures account for approximately 20% of the rating for a Medicare Advantage plan, due to high weightage given to the PQA medication safety and adherence measures.[115] Moreover, data on several other medication safety and adherence measures are reported on a public domain to reinforce quality improvement of Medicare health plans.[115] Despite an abundance of research directed at Medicare plans, there is sparse evidence of the quality of Medicaid programs and few initiatives to address quality measurement of Medicaid programs, particularly for the adult population.

The Affordable Care Act mandated the Secretary of Health and Human Services (HHS) to establish a Medicaid adult quality measurement program by January 2012.[5] In response to the mandate, AHRQ recommended an initial list of 51 measures to assess Medicaid quality for public comment, which included several medication use-related measures. Approximately one-third of the public comments indicated the burden of reporting by states on 51 measures. Other comments alluded to aligning measures with federal measures and ensuring that the measures met the thresholds for evidence, validity, reliability, and feasibility.[116] Based on the public feedback, the measure list was revised to 24 measures including two measures related to long-term management of chronic conditions using medications.

However, this does not undermine the importance of measures included in the study. In fact, the findings presented in this study provide evidence for the importance of inclusion of

these measures in sight of variations in care provided and addresses the issue of feasibility of computing these measures using existing data sources. Though few medication use-related measures are currently being considered for evaluation of Medicaid programs, we believe future quality measurement efforts will include more measures related to medication safety and adherence, as is the case with Medicare.

In this study, Medicaid administrative claims data were demonstrated to be capable of providing state level performance measure scores on all proposed medication use-related quality measures except MI and MI1 measures. The MI measure could be assessed only for 30 states and the MI1 measure could be computed for only 36 states because the remaining states had less than 30 patients eligible for the measure. There was considerable variation in average performance scores across the 13 study measures. The average performance score was lowest at 31.3% for the measure assessing persistence with antidepressant therapy over a six-month period, which is the measure retained in the final Medicaid adult quality measure set. The coefficient of variation was substantially higher for the AD chronic measure compared to all other adherence/persistence measures included in the study. States performed best on the standard of care CAD measure that assessed the use of statins in patients with coronary artery disease with 63.0% receiving the recommended medication. Our study results should motivate future research to focus on the reasons for the variation in quality of care. Identifying Medicaid plans and practitioners providing good medication use-related quality could help to inform other plans on strategies to reduce or eliminate variations in quality of care provided.

It is also important to highlight the room for improvement on the study measures across all states. Direct comparisons with 2012 CMS benchmarks for Medicare Advantage plans reveal that even the best performing states had performance scores lower than the 3-star rated Medicare

Advantage plans. For instance, the 2012 PDC-ACEI/ARB and PDC-STAT thresholds for a 3-star plan are 70.1% and 67.4% respectively, and the corresponding measure scores for the best performing state Medicaid programs in 2007 are 65.2% and 64.0% respectively.[115] The federal government could provide monetary incentives to Medicaid programs with high performance scores on these measures to encourage performance improvement. A bonus payment strategy to reward the top performing states may serve to motivate states to develop interventions targeted at quality improvement as opposed to the cost-containment strategies currently being employed by Medicaid programs. Rather than identifying the top performing plans relative to each other, fixed benchmarks should be set for Medicaid programs considering the need for substantial improvements in medication use-related quality across all states.

States will start reporting on Medicaid quality measures in 2014, which includes the AD acute and AD chronic measures included in our study.[116] This study concentrated only on chronic medication use-related measures. Future research should explore the other 23 measures in the AHRQ final set of Medicaid adult measures. As public reporting of Medicaid quality becomes a reality, it is important to assess if there is a need to adjust the performance of states for patient case-mix.

Our study compares unadjusted performance scores on medication use-related measures with two case-mix adjusted scores based on logistic regression and hierarchical logistic regression models with state as a random intercept. The results from both the classical logistic regression model and the hierarchical logistic regression model indicate that all patient level predictors and the state random effect are significant predictors of the measures. Furthermore both models displayed modest predictive ability (c-statistic > 0.6) for all adherence measures, and good predictive ability for the CAD measure (c-statistic=0.74). The models performed

poorly in case of the ICS measure invalidating the need for case-mix adjustments using the proposed patient characteristics. Future studies should explore the use of additional patient characteristics to improve the model prediction, particularly for the ICS measure. Comparisons of the c-statistics of both models revealed a lack of improvement in discriminative ability of the models when the state-level random effect was added. This finding was reinforced by low residual ICC observed for the state-level across all measures. State level explained small but meaningful amount of variation in the two standard of care measures, 5.7% in case of MI1 measure and 4.1% for the CAD measure. This was further reduced for the PDC measures, with state level explaining approximately between 1.5 – 2% of the variation. These findings suggest that state level factors may not be as important as patient factors in explaining the variation in medication use-related measure scores.

States were ranked on case-mix adjusted scores and unadjusted scores and categorized into top (~20%), medium (60%) and bottom (~20%) performers. States were categorized accordingly to mimic the CMS categorization of hospitals and nursing homes for pay-for-performance demonstration projects.[113,114] Results of the agreement in ranks based on the unadjusted and case-mix adjusted scores showed that case-mix adjustment ranked the states very differently with the majority of states (>50%) changing more than two ranks across all measures except ICS. The validity of the risk adjustment model for the ICS measure was questionable because of the poor discriminative ability of the model. Moreover, there was a lack of agreement in grouping of states into top and bottom performers based on crude and case-mix adjusted methods for majority of the measures (κ=0.22 – 0.74), except for the ICS measure (κ=0.91). The lack of agreement in crude and case-mix adjusted scores was also observed in case of the composite measure (κ=0.65). Mehta et al (2008) assessed the agreement in grouping of hospitals

based on unadjusted and risk adjusted rankings of a composite measure score and observed similar changes in hospital performance rankings.[80]

The two risk adjustment methods used in the study showed high level of agreement across all measures. For the composite measure, there was perfect agreement in groups based on the two-level and the three-level hierarchical regression models (κ=1.00). Overall, these findings reinforce the importance of case-mix adjustment using either classical logistic regression models or hierarchical logistic regression models in providing performance scores for state Medicaid programs. The importance of case-mix adjusting of performance scores for public reporting or pay-for-performance purposes has been reported in several studies. Several researchers have found that adjusting for patient demographics, disease severity and comorbidity burden changes providers', facilities' or healthcare systems' performance profiles.[78-83] Hofer et al.[6] demonstrated that physicians can improve their performance profile by preferential patient selection. Similarly, a recent study found that adjusting for patient characteristics and treatment opportunities improved hospital rankings on indicators assessing adherence to treatment guidelines for acute MI.[80] The importance of case-mix adjustment was also demonstrated in evaluating health plan performance on chronic care provided to Medicaid enrollees.[82] Similarly, this study emphasizes the importance of risk-adjustment methodologies to adjust for patient differences in assessing the quality provided by state Medicaid programs.

Study findings also demonstrate that racial disparities predominantly drive the need for case-mix adjustment of medication-use related quality measures. Blacks were significantly less likely to receive good quality of care (measured as adherence, persistence with therapy and receiving standard of care medications) compared to Whites on 11 of the 13 measures studied, after controlling for demographic characteristics and comorbidity burden. Hispanics were also

less likely to receive standard of care medications, and be adherent/persistent to the recommended therapy compared to Whites. On the contrary, Blacks and Hispanics were more likely to fill an inhaled corticosteroid or similar medication to manage persistent asthma compared to Whites. The racial disparities were substantial and consistent across the measures included in the study. However, the magnitude of the disparity varies across the quality measures. The disparity was smallest for the MI1 measure and largest for the AD measures. This study's findings reinforce the existing evidence of disparities in adherence and persistence with medication therapy.[40,67,68]

This study did not control for socioeconomic characteristics like education and income. Despite the population being Medicaid enrollees, states have different enrollment criteria and socioeconomic variables may be useful in explaining some of the racial disparity in quality of care observed. In addition, future studies should explore the effect of provider and payer characteristics in quality of care provided to different racial groups. Racial minorities may be disproportionately enrolled in poorly performing plans and receiving care from low quality providers. Medicaid programs need to identify the quality of care being provided by various practitioners and payers serving Medicaid beneficiaries and incentivize top performers to provide better quality of care.

This widespread disparity in the medication use-related quality of care received by different racial/ethnic groups is formidable. Health plans and practitioners have an important role to play in improvement of medication utilization patterns among Blacks and Hispanics. Existing evidence emphasizes the positive impact of practitioner-led initiatives in closing the racial disparities in care provision.[117] Additionally, payers are taking various medication therapy management initiatives to monitor and improve patient safety and adherence with medication

regimen. Medicaid data is a useful data source to identify underserved populations, and Medicaid plans can use the results of this study to provide targeted medication management interventions to racial minority beneficiaries.

An objective of this study was to assess if states performed consistently across all the medication use-related measures. While there was some degree of correlation between the state rankings on the seven adherence measures, states were ranked differently on the standard of care measures. Considering the substantial differences in performance of states on measures related to the management of chronic conditions using medication therapy, it becomes crucial to construct one unique measure of medication quality for each state that can be used for public reporting and pay-for-performance.

A composite measure of chronic medication use-related quality was constructed adjusting both for patient case-mix and opportunity mix. The average composite score across states was 53.1%, with values ranging from 46.0% for Mississippi to 63.2% for Vermont. Though some states are performing relatively poorer than others, all Medicaid programs demonstrated a need for improvement on medication use-related measures. Medicaid programs can achieve quality improvements by diverting its patients to better quality providers in the state and by incentivizing providers providing good quality care. Medicaid can also employ quality bonus payments similar to Medicare to motivate providers to improve performance. However, the budget shortfalls that most states are experiencing present challenges in adopting a pay-for-performance strategy. Therefore, before financial incentives are considered, public reporting of performance is recommended.

The results of this study should be viewed in the light of a few limitations. Inherent limitations of administrative claims data research apply to this study. It is not possible to assess the reasons for discontinuation of medications using administrative claims data. For example, change in therapy by the physician due to lack of efficacy or adequate control through alternative methods may be the reason for discontinuation of medications. Another limitation of using administrative claims database is that, factors explaining adherence and persistence measures like self-efficacy, patient-provider communication, health literacy, or readiness to change cannot be measured using claims data. Also, it is not possible to capture medication refills not paid for by Medicaid using Medicaid data. These inadequacies in administrative claims data may have biased our study results. Another limitation is the inability to concretely judge the need for risk adjustment methods. Since we do not know the true rankings of quality, it is challenging to ascertain rankings based on one method as being more "right" than the other. However, since there is evidence of difference in patient case-mix across states and because these patient characteristics are significantly associated with medication utilization patterns, one can postulate that case-mix adjusted scores are more appropriate indicators of Medicaid quality than the unadjusted scores.

This study has several strengths. To our knowledge, this is the first study to examine interstate variations in quality of care provided using a comprehensive list of medication use-related measures spanning several chronic conditions. This study provides evidence for the feasibility of assessing medication use-related quality measures for Medicaid programs using administrative claims data. This study highlights the need for including medication use-related measures in the Medicaid adult quality measure set, considering the variations in performance of states on these measures and the substantial room for improvement across all measures. National

benchmarks on medication use-related measures for the year 2007 are computed. Eventually, when states start voluntarily reporting on medication use measures, the study findings provide a baseline against which to compare changes in medication use-related quality. Additionally, I created a composite measure of medication use-related quality adjusting for patient case-mix and opportunity-mix. States could use a similar approach to report on medication use-related quality if they find it too burdensome to report on multiple measures. Moreover, the importance of case-mix adjustment of performance scores when comparing state Medicaid programs is demonstrated. The study showcased the lack of agreement in rankings and reimbursement grouping of states based on crude and case-mix adjusted scores. Medicaid programs should consider the findings of this study before publicly reporting on crude performance scores.

LIST OF REFERENCES

1. Department of Health & Human Services. Hospital compare. Available at www.hospitalcompare.hhs.gov/Hospital/Search/Welcome.asp. Accessed on April 7, 2011.
2. Department of Health & Human Services. Nursing home compare. Available at www.medicare.gov/NHcompare. Accessed on April 7, 2011.
3. National Committee for Quality Assurance. Health plan report card. Available at http://reportcard.ncqa.org/plan/external/plansearch.aspx. Accessed on April 7, 2011.
4. Rosenthal MB, Landrum MB, Meara E, Huskamp HA, Conti RM, Keating NL. Using performance data to identify preferred hospitals. *Health Serv Res.* 2007;42:2109-2119.
5. Department of Health & Human Services. Medicaid program: initial core set of health quality measures for Medicaid-eligible adults. Available at http://www.gpo.gov/fdsys/pkg/FR-2010-12-30/pdf/2010-32978.pdf. Accessed on April 7, 2010.
6. Hofer TP, Hayward RA, Greenfield S, Wagner EH, Kaplan SH, Manning WG. The unreliability of individual physician "report cards" for assessing the costs and quality of care of a chronic disease. *JAMA*. 1999;281:2098-2105.
7. Kerr EA, Krein SL, Vijan S, Hofer TP, Hayward RA. Avoiding pitfalls in chronic disease quality measurement: a case for the next generation technical quality measures. *Am J Manag Care*. 2001;7:1033-1043.
8. Jayasinghe UW, Proudfoot J, Barton CA, et al. Quality of life of Australian chronically-ill adults: patient and practice characteristics matter. *Health Qual Life Outcomes.* 2009;7:50-61.
9. Si D, Bailie R, Dowden M, et al. Assessing quality of diabetes care and its variation in Aboriginal community health centers in Australia. *Diabetes Metab Res Rev*. 2010;26:464-473.
10. Pharmacy Quality Alliance. PQA approved measures. Available at http://www.pqaalliance.org/measures.htm. Accessed on April 7, 2010.
11. Pharmacy Quality Alliance. PQA draft measures for testing in 2010. Available at http://www.pqaalliance.org/measures.htm. Accessed on April 7, 2010.
12. Cunningham PJ. State variation in primary care physician supply: implications for health reform Medicaid expansions. Research Brief No. 19. Washington, DC: Center for Studying Health System Change; 2011.
13. Coughlin TA, Long SK. Adult health care access and use under Medicaid: does it vary by state? *J Health Care Poor Underserved.* 2003;14:208-228.
14. National Committee for Quality Assurance. HEDIS 2011 measures. Available at http://www.ncqa.org/tabid/1223/Default.aspx. Accessed on April 7, 2011.
15. American Medical Association. Physician Consortium for Performance Improvement measures. Available at www.ama-assn.org. Accessed on April 7, 2011.
16. Nolte E, McKee CM. Measuring the health of nations: updating an earlier analysis. *Health Aff.* 2008;27:58-71.
17. Nau DP. Measuring pharmacy quality. *J Am Pharm Assoc*. 2009;49:154-163.
18. Rosenthal MB, Fernandopulle, Song HR, Landon B. Paying for quality: providers' incentives for quality improvement. *Health Aff.* 2004;23:127-141.

19. Reschovsky J, Hadley J. Physician financial incentives: use of quality incentives inches up, but productivity still dominates. Issue Brief no.108. Washington DC: Center for Studying Health system Change; 2007.
20. Draper DA, Liebhaber A, Ginsburg PB. High-performance health plan networks: early experiences. Issue Brief no. 111. Washington, DC: Center for Studying Health System Change; 2007.
21. Dudley RA, Miller RH, Korenbrot TY, et al. The impact of financial incentives on quality of healthcare. *Milbank q*. 1998;76:649-686.
22. Institute of Medicine. Crossing the Quality Chasm: The IOM Health Care Quality Initiative. Available from: http://www.iom.edu/Reports/2001/. Accessed on 31 March 2011.
23. Donabedian A. The quality of care. How can it be assessed? *JAMA.* 1988;260:1743-1748.
24. Kozma CM, Reeder CE, Schulz RM. Economic, clinical and humanistic outcomes: a planning tool for pharmacoeconomic research. *Cin Ther.* 1993;15:1121-1132.
25. Campbell SM, Braspenning J, Hutchinson A, Marshall M. Research methods used in developing and applying quality indicators in primary care. *Qual Saf Health Care*. 2002;11:358-364.
26. Iezzoni LI. Assessing quality using administrative data. *Ann Intern Med.* 1997;127:666-674.
27. Pillittere-Dugan D, Nau DP, McDonough K, Zakiya P. Development and testing of performance measures for pharmacy services. *J Am Pharm Assoc.* 2009;49:212-219.
28. Winsdale N, Taylor L, Shi S, Schuwirth L, Vleuten CV, Tamblyn R. Monitoring community pharmacists's quality of care: a feasibility study of using pharmacy claims data to assess performance. *BMC Health Serv Res*. 2011;11:12.
29. Eddy DM. Performance measurement: problems and solutions. *Health Aff.* 1998;17:7-25.
30. Brook RH, McGlynn EA, Cleary PD. Quality of health care, part 2: measuring quality of care. *N Engl J Med.* 1996;335:966-970.
31. Yusuf S, Wittes J, Friedman L. Overview of results of randomized clinical trials in heart disease: I. Treatments following myocardial infarction. *JAMA.* 1988;260:2088-2093.
32. Freemantle N, Cleland J, Young P, et al. Beta blockade after myocardial infarction: systematic review and meta regression analysis. *BMJ*. 1999;318:1730-1737.
33. Dargie HJ, et al, for the CAPRICORN Investigators. Effect of carvedilol on outcome after myocardial infarction in patients with left-ventricular dysfunction: the CAPRICORN randomised trial. *Lancet.* 2001;357:1385-390.
34. Krumholz HM, Radford MJ, Wang Y, et al. National use and effectiveness of beta-blockers for the treatment of elderly patients after acute myocardial infarction: National Cooperative Cardiovascular Project. *JAMA.* 1998;280:623-629.
35. Bradford WD, Chen J, Krumholz HM. Under-utilisation of betablockers after acute myocardial infarction. Pharmacoeconomic implications. *Pharmacoeconomics*. 1999;15:257-268.
36. Kramer JM, Hammill B, Anstrom KJ, Fetterolf D, Snyder R, Charde JP et al. National evaluation of adherence to beta-blocker therapy for 1 year after acute myocardial infarction in patients with commercial health insurance. *Am Heart J.* 2006;152(3):454.e1-454.e8.
37. National Committee for Quality Assurance. The State of Health Care Quality: 2003. Available at http://www.ncqa.org/communications/State%20Of%20Managed%20Care/SOHCREPORT2003.pdf. Accessed on April 7, 2011.

38. Eagle KA, Kline-Rogers E, Goodman SG, et al. Adherence to evidence-based therapies after discharge for acute coronary syndromes: an ongoing prospective, observational study. *Am J Med.* 2004;117:73–81.
39. Ackincigil A, Bowblis JR, Levin C, Jan S, Patel M, Crystal S. Long-term adherence to evidence-based secondary prevention therapies after acute myocardial infarction. *J Gen Intern Med.* 2008;23(2):115-121.
40. Choudhry NK, Setoguchi S, Levin R, Winkelmayer WC, Shrank WH. Trends in adherence to secondary prevention medications in elderly post-myocardial infarction patients. *Pharmacoepidemiol Drug Saf.* 2008;17(12):1189-1196.
41. Shepherd J, Cobbe SM, Ford I, Isles CG, Lorimer AR, MacFarlane PW, et al. Prevention of coronary heart disease with pravastatin in men with hypercholesterolemia. West of Scotland Coronary Prevention Study Group. *N Engl J Med*. 1995;333:1301–7.
42. Randomised trial of cholesterol lowering in 4444 patients with coronary heart disease: the Scandinavian Simvastatin Survival Study (4S). *Lancet*. 1994;344:1383–9.
43. Prevention of cardiovascular events and death with pravastatin in patients with coronary heart disease and a broad range of initial cholesterol levels. The Long-Term Intervention with Pravastatin in Ischaemic Disease (LIPID) Study Group. *N Engl J Med*. 1998;339:1349–57.
44. Tonkin A, Aylward P, Colquhoun D, Glasziou P, et al. Prevention of cardiovascular events and death with pravastatin in patients with coronary heart disease and a broad range of initial cholesterol levels. *New Engl J Med*. 1998; 339:1349-1358.
45. National Cholesterol Education Program. Detection, Evaluation, and Treatment of High Blood Cholesterol in Adults (Adult Treatment Panel III).2002. Available at http://www.nhlbi.nih.gov/guidelines/cholesterol/index.htm. Accessed on April 4, 2011.
46. EUROASPIRE. A European Society of Cardiology survey of secondary prevention of coronary heart disease: principal results. EUROASPIRE Study Group. European Action on Secondary Prevention through Intervention to Reduce Events. *Eur Heart J*. 1997;18:1569–1582.
47. Whincup PH, Emberson JR, Lennon L, Walker M, Papacosta O, et al. Low prevalence of lipid lowering drug use in older men with established coronary heart disease. Heart. 2002;88:25–29.
48. Sueta CA, Chowdhury M, Boccuzzi SJ, Smith SC Jr, Alexander CM, Londhe A, Lulla A, Simpson RJ Jr. Analysis of the degree of undertreatment of hyperlipidemia and congestive heart failure secondary to coronary artery disease. *Am J Cardiol.* 1999;83:1303–1307.
49. Education and Prevention Program's Expert Panel Report 3 (EPR 3) on the diagnosis and management of asthma. Full report.2007. http://www.nhlbi.nih.gov/guidelines/asthma/asthgdln.pdf. Accessed April 4,2011.
50. Suissa S, Ernst P. inhaled corticosteroids: impact on asthma morbidity and mortality. *J Allergy Clin Immunol.* 2001;107:937-944.
51. Australian Centre for Asthma Monitoring (2007). Asthma in Australia: findings from the 2004–05 National Health Survey. AIHW cat. no. ACM 10. Canberra: Australian Institute of Health and Welfare. Available at: http://www.asthmamonitoring.org/publications.htm. Accessed April 4,2011.
52. Kandane-Rathnayake1 RK, Matheson MC, Simpson JA, et al. Adherence to asthma management guidelines by middle-aged adults with current asthma. *Thorax*.2009;64:1025-1031.

53. American Psychiatric Association: Practice guideline for the treatment of patients with major depressive disorder (Revision). *Am J Psychiatry*. 2000;157.
54. Agency for Health Care Policy Research: Treatment of Major Depression: Clinical Practice Guidelines 5, vol 2. Rockville, MD, US Department of Health and Human Services, 1993.
55. Schulberg HC, Katon W, and Simon GE. et al. Treating major depression in primary care practice: an update of the Agency for Health Care Policy and Research practice guidelines. *Arch Gen Psychiatry*. 1998. 55:1121–1127.
56. National Institute for Clinical Excellence. Depression: the treatment and management of depression in adults (update). Available at: www.nice.org.uk. Accessed April 02, 2011.
57. Geddes JR, Carney SM, Davies C, et al. Relapse prevention with antidepressant drug treatment in depressive disorders. *Lancet*. 2003;361:653–651.
58. Lin EH, Katon WJ, VonKorff M, et al. Relapse of depression in primary care: rate and clinical predictors. *Arch Fam Med*. 1998;7:443–449.
59. Keller MB, Hirschfeld RMA, Kemyyenaere K, Baldwin DS. Optimizing outcomes in depression: focus on antidepressant compliance. *Int Clin Psychopharmacol*. 2002;17:265–271.
60. Fairman KA, Drevets WC, and Kreisman JJ. et al. Course of antidepressant treatment drug type, and prescriber's specialty. *Psychiatr Serv.* 1998;49:1180–1186.
61. Olfson M, Marcus SC, and Tedeschi M. et al. Continuity of antidepressant treatment for adults with depression in the United States. *Am J Psychiatry*. 2006;163:101–108.
62. Bull SA, Hunkeler EM, and Lee JY. et al. Discontinuing or switching selective serotonin-reuptake inhibitors. *Ann Pharmacother.* 2002;36:578–584.
63. Dunn RL, Donoghue JM, and Ozminkowski RJ. et al. Selective serotonin reuptake inhibitor antidepressant prescribing in primary care in the United Kingdom: a longitudinal analysis. *Prim Care Psychiatry.* 1998;4:141–148.
64. Melfi CA, Chawla AJ, Croghan TW, Hanna MP, Kennedy S, Sredl K. The effects of adherence to antidepressant treatment guidelines on relapse and recurrence of depression. *Arch Gen Psychiatry*. 1998; 55:1128–1132.
65. DiMatteo MRP: Variations in patients' adherence to medical recommendations: a quantitative review of 50 years of research. *Med Care*. 2004; 42: 200-209.74
66. Hansen R, Seifeldin R, Noe L. Medication Adherence to Chronic Disease: Issues in Posttransplant Immunosupression. *Transplantation Proceedings*. 2009;39:1287-1300.
67. Cramer JA: A systematic review of adherence with medications for diabetes. *Diabetes Care*. 2004; 27:1218-1224.
68. Yang Y, Thumula V, Pace PF, et al. Medication nonadherence and the risks of hospitalization, emergency department visits, and death among Medicare part D enrollees with diabetes. *Drug Benefit Trends*. 2009;21:12.
69. Simons LA, Levis G, Simons J. Apparent discontinuation rates in patients prescribed lipid-lowering drugs. *Med J Aust.* 1996;164:208-211.
70. Gerth WC. Compliance and persistence with newer antihypertensive agents. *Curr Hypertens Rep*. 2002;4:424-433.
71. Schectman JL, Nadkarni MM, Voss JD. The association between diabetes metabolic control and drug adherence in an indigent population. *Diab Care.* 2002;25:1015-1021.

72. Pignone M, Phillips C, Mulrow C. Use of lipid lowering drugs for primary prevention of coronary heart disease: meta-analysis of randomised trials. *BMJ*. 2000;321:983-986.
73. Ross SD, Allen IE, Connelly JE, et al. Clinical outcomes in statin treatment trials: a meta-analysis. *Arch Intern Med.* 1999;159:1793-1802.
74. Neal B, MacMahon S, Chapman N. Blood Pressure Lowering Treatment Trialists' Collaboration. Effects of ACE inhibitors, calcium antagonists, and other bloodpressure–lowering drugs: results of prospectively designed overviews of randomised trials. *Lancet.* 2000;356:1955-1964.
75. Law MR, Wald NJ, Morris JK, Jordan RE. Value of low dose combination treatment with blood pressure lowering drugs: analysis of 354 randomised trials. *BMJ.*2003;326:1427-1431.
76. ALLHAT Officers and Coordinators for the ALLHAT Collaborative Research Group. Major outcomes in moderately hypercholesterolemic, hypertensive patients randomized to pravastatin vs usual care: the Antihypertensive and Lipid-Lowering Treatment to Prevent Heart Attack Trial (ALLHAT-LLT). *JAMA.* 2002;288:2998-3007.
77. Resnik L, Liu D, Hart DL, Mor V. Benchmarking physical therapy clinic performance: statistical methods to enhance internal validity when using observational data. *Phys Ther*. 2008;88:1078-1087.
78. Greenfield S, Kaplan SH, Kahn R, Ninomiya J, Griffith JL. Profiling case provided by different groups of physicians: effects of patient case-mix (bias) and physician-level clustering on quality assessment results. *Ann Intern Med.* 2002;136:111-121.
79. Rosen AK, Reid R, Broemeling A, Rakovski CC. Applying risk-adjustment framework to primary care: can we improve on existing measures? *Ann Fam Med.* 2003;1:44-51.
80. Mehta RH, Liang L, Karve AM, et al. Association of patient case-mix adjustment, hospital process performance rankings, and eligibility for financial incentives. *JAMA.* 2008;300:1897-1903.
81. Liu C, Sales AE, Sharp ND, et al. Case-mix adjusting performance measures in a veteran population: pharmacy- and diagnosis-based approaches. *Health Ser Res.* 2003;38:1319-1337.
82. Volpel A, O'Brien J, Weiner J. Strategies for assessing health plan performance on chronic diseases: selecting performance indicators and applying health based risk adjustment. 2005. Available at http://www.chcs.org/usr_doc/99529MD.pdf. Accessed on April 7, 2011.
83. Escobar GJ, Greene JD, Scheirer P, Gardner MN, Draper D, Kipnis P. Risk-adjusting hospital inpatient mortality using automated inpatient, outpatient, and laboratory databases. *Med Care.* 2008;46:232-239.
84. Weiner JP, Starfield BH, Steinwachs DM, et al. Development and application of a population-oriented measure of ambulatory case-mix. *Med Care.* 1991;29:452–472.
85. Kronick R, Gilmer T, Dreyfus T, Lee L. Improving health-based payment for Medicaid beneficiaries: The Chronic and Disability Payment System. *Health Care Finan Rev.* 2000;21:29–64.
86. Fishman PA, Goodman MJ, Hornbrook MC, Meenan RT, Bachman DJ, O'Keeffe Rosetti MC. Risk adjustment using automated ambulatory pharmacy data: the RxRisk model. Med Care. 2003;41:84-99.
87. Gilmer T, Kronick R, Fishman P, et al. The Medicaid RX model: Pharmacy-based risk adjustment for public programs. *Med Care.* 2001;39:1188–1202.
88. Ellis RP, Pope GC, Iezzoni L, et al. Diagnosis-based risk adjustment for Medicare capitation payments. *Health Care Financ Rev.* 1996;17:101-128.
89. Clark DO, Von Korff M, Saunders K, et al. A chronic disease score with empirically derived weights. *Med Care.* 1995;33:783–795.

90. Lamers LM. Pharmacy costs groups: a risk-adjuster for capitation payments based on the use of prescribed drugs. *Med Care.* 1999;37:824–830.
91. Fishman P, Shay D. Development and estimation of a pediatric chronic disease score from automated pharmacy data. *Med Care.* 1999;37:872–880.
92. Ou H. Performances of comorbidity measures in healthcare related behaviors and outcomes in type 2 diabetes. 2010. Available at http://hdl.handle.net/2027.42/77787. Accessed on April 7, 2011.
93. Lu CY, Barratt J, Vitry A, Roughead E. Charlson and Rx-Risk comorbidity indices were predictive of mortality in the Australian health care setting. J Clin Epidemiol. 2011;64:223-228.
94. Yang Y, Thumula V, Pace PF, et al. Nonadherence to angiotensin converting enzyme inhibitors and/or angiotensin II receptor blockers among high-risk patients with diabetes in Medicare part D programs. *J Am Pharm Assoc.* 2010;50:527-531.
95. Wu SY, Green A. Projection of chronic illness prevalence and cost inflation. RAND Corporation. 2000.
96. Center for Disease Control and Prevention. The power to prevent, the call to control: at a glance 2009. 2009. Available at http://www.cdc.gov/chronicdisease/resources/ . Accessed on April 11, 2010.
97. Rand CS,Wise RA. Measuring adherence in asthma medication regimens. *Am J Resp Crit Care Med*. 1994;149:S69–76.
98. Arday DR, Fleming BB, Keller DK, et al. Variation in diabetes care among states: do patient characteristics matter? *Diabetes Care.* 2002;25:2230-2237.
99. Gulliford MC, Mahabir D, Ukoumunne OC. Evaluating variations in medical practice between government primary care health centers. *J Clin Epidemiol.* 2001;54:511-517.
100. Krein SL, Hofer TP, Kerr EA, Hayward RA. Whom should we profile? Examining diabetes care practice variation among primary care providers, provider groups, and health care facilities. *Health Serv Res*. 2002;37:1159-1180.
101. O'Conner PJ, Rush WA, Davidson G, et al. Variation in quality of diabetes care at the levels of patient, physician, and clinic. *Prev Chronic Dis*. 2008;5. Available at http://www.cdc.gov/pcd/issues/2008/jan/06_0118.htm. Accessed on April 7, 2011.
102. Fung V, Schmittdiel JA, Fireman B, et al. Meaningful variation in performance: a systematic literature review. *Med Care*. 2010;48:140-148.
103. Aday LA, Andersen RM. A framework for the study of access to medical care. *Health Serv Res*. 1974;9:208-220.
104. Tang MH, Hill KS, Boudreau AA, Yucel RM, Perrin JM, Kuhlthau KA. Medicaid managed care and the unmet need for mental health care among children with special health care needs. *Health Serv Res*. 2008;43:882-900.
105. Ward E, Halpern M, Schrag N, et al. Association of insurance with cancer care utilization and outcomes. *CA Cancer J Clin.* 2008;58:9–31.
106. ResDAC. Medicaid data available. Available at http://www.resdac.org/Medicaid/data_available.asp. Accessed on April 7, 2010.
107. Peterson AM, Nau DP, Cramer JA, Benner J et al. A checklist for medication compliance and persistence studies using retrospective databases. *Value Health*.2007;10(1):3-12.
108. Ho PM, Rumsfeld JS, Masoudi FA, et al. Effect of medication nonadherence on hospitalization and mortality among patients with diabetes mellitus. *Arch Intern Med*. 2006;166:1836–1841.

109. Johnson ML, El-Serag HB, Tran TT, Hartman C, Richardson P, Abraham NS. Adapting the Rx-Risk-V for mortality prediction in outpatient populations. *Med Care.* 2006;44:793–797.

110. SAS® software, version 9.3.SAS Institue, Cary, NC,USA.

111. Hox JJ. Multilevel Analysis: Techniques and Applications. 2nd ed. New York, NY: Routledge;2010.

112. Snijders TAB, Bosker RJ. Multilevel analysis: An introduction to basic and advanced multilevel modeling. London, England:Sage;1999.

113. Li Y, Cai X, Glance LG, Spector WD, Mukamel DB. National release of the nursing home quality report cards: implications of statistical methodology for risk adjustment. *Health Serv Res*. 2009;44(1):79–102.

114. Lindenauer PK, Remus D, Roman S, et al. Public reporting and pay-for-performance in hospital quality improvement. *New Eng J Med*. 2007; 356(5):486−496.

115. Executive update on Medication Quality Measures in Medicare Part D Star Ratings. Available at http://www.pqaalliance.org/files/2012UpdateOnMedicarePartDStarRatings.pdf. Accessed on May 27, 2012.

116. Department of Health & Human Services. Medicaid program: initial core set of health quality measures for Medicaid-eligible adults. Available at http://federalregister.gov/a/2011-33756. Accessed on April 7, 2010.

117. Sequist TD, Adams A, Zhang F, Ross-Dengan D, Ayanian JZ. Effect of quality improvement on racial disparities in diabetes care. *Arch Intern Med*. 2006;166(6):675-681

LIST OF APPENDICES

APPENDIX-A: CONCEPTUAL FRAMEWORK

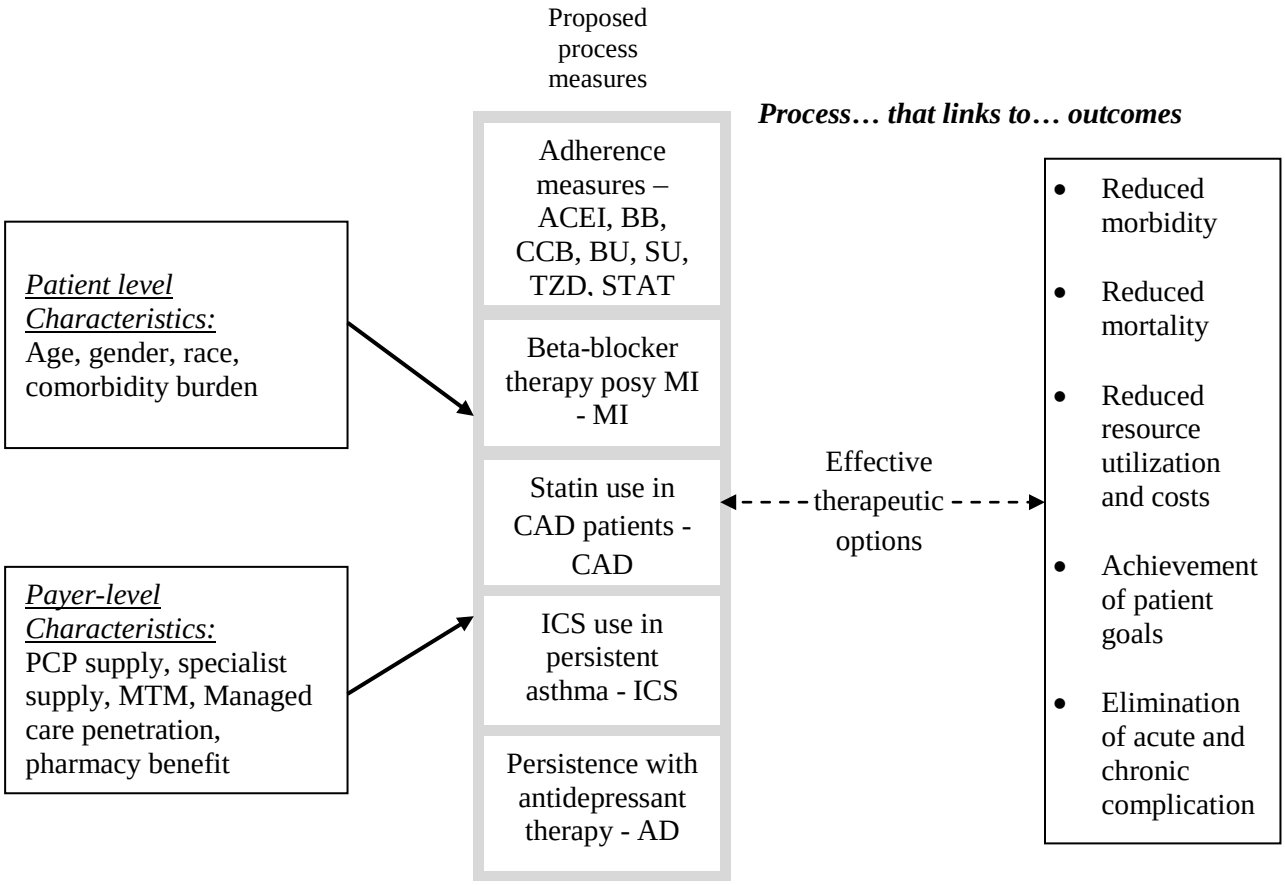

FIGURE A-1: Conceptual Framework Underlying the Study.

APPENDIX-B: ADAY-ANDERSON MODEL

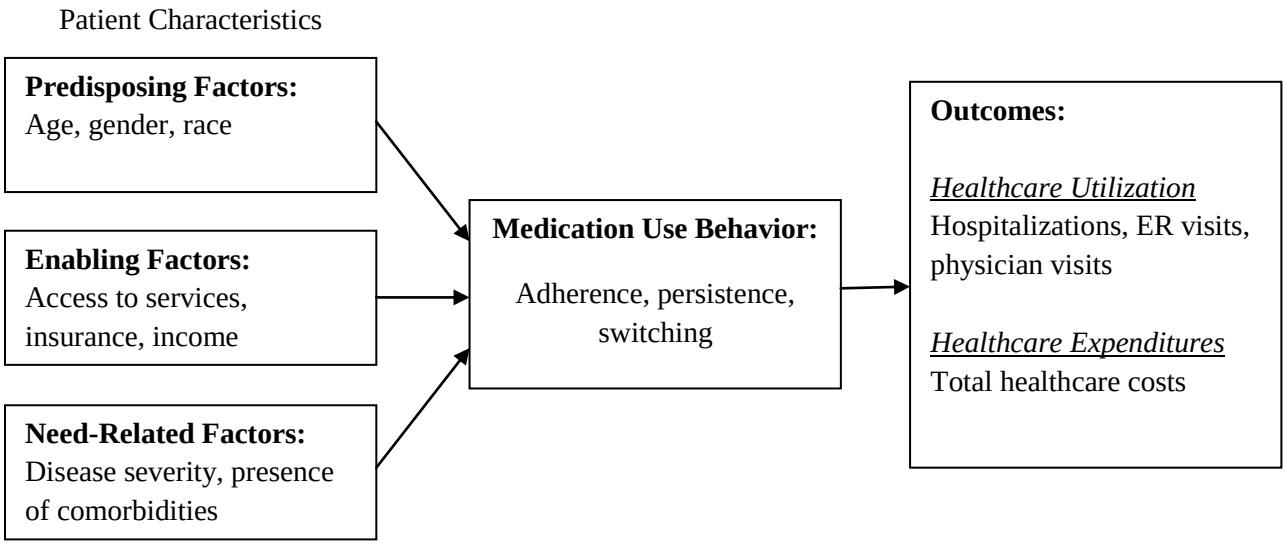

FIGURE A-2: The Aday-Andersen Model for Determinants of Healthcare Utilization[33]

APPENDIX-C: DEFINITIONS OF STUDY MEASURES

TABLE A-1: Definitions of Study Measures

Indicator	Definition
MI – Persistence of beta-blocker treatment after a heart attack	*Measure:* Proportion of patients who had a myocardial infarction that are persistent to beta-blocker therapy *Numerator:* # patients with continuous (gap of less than 51 days or less) use of beta-blockers for at least 6 months during the measurement period. *Denominator:* # of patient who had a myocardial infarction between April 1, 2006 and March 31, 2007 and filled a beta-blocker prescription within 30 days of discharge.
CAD – Statin therapy in patients with coronary artery disease	*Measure:* Proportion of patients with coronary artery disease who received at least one prescription for a statin. *Numerator:* # patients with at least one prescription fill for a statin medication during the measurement period. *Denominator:* # patients with a diagnosis for coronary artery disease during the measurement year.
ICS – Use of appropriate medications for people with asthma	*Measure:* Proportion of patients with persistent asthma who were dispensed an inhaled corticosteroid or similar medication during the measurement period. *Numerator:* # patients with at least one prescription for inhaled corticosteroid, nedocromil, cromolyn sodium, leukotriene modifiers or methylxanthines during the measurement period. *Denominator:* # patients with persistent asthma, ages 18-56 years.
AD – Persistence with antidepressant medications	*Measure:* Proportion of patients on antidepressant therapy that are persistent with therapy for at least 6 months. *Numerator:* # patients with continuous use of antidepressants for at least 6 months during the measurement period (total gap of 51 days or less). *Denominator:* # patients with at least two prescriptions for an antidepressant medication filled on two unique service dates, with the first prescription filled between May 1, 2006 and April 30, 2007 and no prior use of antidepressants for at least three months.
PDC - Adherence with medications	*Measure:* Proportion of patients who met the PDC threshold of 80% for

	Beta-blockers (BB), ACEI/ARBs (ACEI), Calcium channel blockers (CCB), Statins (LLD), Biguanides (BU), Sulfonylureas (SU), Thiazolidinedione (TZD). *Numerator:* # patients who met the 80% PDC threshold *Denominator:* # patients with at least two prescription fills for medications in the particular therapeutic class during the measurement period.

- Additionally, patients should be of age 18-<65 years, except for asthma, and continuously enrolled in Medicaid during the measurement year.
- Patients with a non acute stay during the measurement year will be excluded for all measures.
- Dual-eligible patients will be excluded.

APPENDIX-D: LIST OF MEDICATIONS

TABLE A-2: List of medications used in the study

ICS MEASURE - ASTHMA – A

Short-Acting Inhaled Beta Agonists

- Albuterol
- levalbuterol
- pirbuterol

Long-Acting Inhaled Beta Agonists and Combinations

- salmeterol
- formoterol
- budesonide & formoterol
- fluticasone & salmeterol

Inhaled Corticosteroids and Combinations

- beclomethasone
- budesonide
- flunisolide
- fluticasone
- fluticasone & salmeterol
- mometasone
- triamcinolone
- budesonide & formoterol

Leukotriene Inhibitors

- zafirlukast
- montelukast
- zileuton

Xanthines

- long acting theophylline

Mast Cell Stabilizers

- nedocromil
- cromolyn

ICS MEASURE - ASTHMA – B

- tiotropium
- ipratropium & albuterol
- ipratropium
- pulmozyme
- beclomethasone
- budesonide
- ciclesonide
- flunisolide
- fluticasone
- mometasone
- triamcinolone

BB, MI, MI1 MEASURES

Beta-Blocker Medications

- acebutolol HCL
- atenolol
- betaxolol HCL
- bisoprolol fumarate
- carteolol HCL
- carvedilol
- labetalol HCL
- metoprolol succinate
- metoprolol tartrate
- nadolol
- nebivolol HCL
- penbutolol sulfate
- pindolol
- propranolol HCL
- timolol maleate

BB Combination Products

- atenolol & chlorthalidone
- bisoprolol & HCTZ
- nadolol & bendroflumethiazide
- metoprolol & HCTZ
- propranolol & HCTZ
- timolol & HCTZ

ACEI/ARB MEASURE

ARB Medications

- candesartan
- eprosartan
- irbesartan
- losartan
- olmesartan
- telmisartan
- valsartan

ACE Inhibitor Medications

- benazepril
- captopril
- enalapril
- fosinopril
- lisinopril
- moexipril
- perindopril
- quinapril
- ramipril
- trandolopril

ACE Inhibitor Combination Products

- amlodipine & benazepril
- benazepril & HCTZ
- captopril & HCTZ
- enalapril & HCTZ
- enalapril & felodipine
- fosinopril & HCTZ
- lisinopril & HCTZ
- moexipril & HCTZ
- lisinopril & nutritional supplement
- quinapril & HCTZ
- trandolopril-verapamil HCL

ARB Combination Products

- candesartan & HCTZ
- eprosartan & HCTZ
- telmisartan & amlodipine
- irbesartan & HCTZ
- losartan & HCTZ
- amlodipine & olmesartan
- olmesartan & HCTZ
- telmisartan & HCTZ
- aliskiren & valsartan
- valsartan & HCTZ
- amlodipine & valsartan
- amlodipine & valsartan & HCTZ

CCB MEASURE

CCB Medications

- amlodipine besylate
- diltiazem HCL
- felodipine
- isradipine
- nicardipine HCL
- nifedipine (long acting only)
- verapamil HCL
- nisoldipine

CCB Combination Products

- amlodipine besylate & benazepril HCL
- amlodipine & valsartan
- amlodipine & valsartan & HCTZ
- enalapril maleate & felodipine
- telmisartan & amlodipine
- amlodipine & olmesartan
- trandolopril & verapamil HCL
- amlodipine & atorvastatin

BIGU MEASURE

Biguanides

- metformin

Biguanide Combination Products

- glipizide & metformin
- glyburide & metformin
- rosiglitazone & metformin
- pioglitazone & metformin
- repaglinide & metformin
- sitagliptin & metformin

SU MEASURE

Sulfonylureas

- acetohexamide
- chlorpropamide
- glimepiride
- glipizide
- glyburide
- tolazamide
- tolbutamide

Sulfonylurea Combination Products

- glipizide & metformin
- glyburide & metformin
- rosiglitazone & glimepiride
- pioglitazone & glimepiride

TZD MEASURE

Thiazolidinediones

- pioglitazone
- rosiglitazone

Thiazolinedione Combination Products

- rosiglitazone & metformin
- pioglitazone & metformin
- rosiglitazone & glimepiride
- pioglitazone & glimepiride

STAT & CAD MEASURE

Statin Medications

- lovastatin
- rosuvastatin
- fluvastatin
- atorvastatin
- pravastatin
- simvastatin

Statin Combination Products

- niacin & lovastatin
- atorvastatin & amlodipine
- niacin & simvastatin
- pravastatin & aspirin
- ezetimibe & simvastatin

APPENDIX-E: REVENUE CODES FOR NON-ACUTE STAYS

TABLE A-3: List of UB Revenue Codes for Identifying Non Acute Stays

Description	UB Revenue
Hospice	0115, 0125, 0135, 0145, 0155, 0650, 0656, 0658, 0659
SNF	019x
Rehabilitation	0118, 0128, 0138, 0148, 0158
Respite	0655
Residential substance abuse treatment facility	1002
Psychiatric residential treatment center	1001

APPENDIX-F: RXRISK CATEGORIES

TABLE A-4: RxRisk Score Categories

Chronic Disease	Medication Class(es)
Anxiety and tension	Salicylate combinations, Barbiturates, Benzodiazepines, Meprobamate, Miscellaneous hypnotics, Paraldehyde
Asthma	Anti-inflammatory glucocorticoids, Isoproterenol, Bronchodilators, Cromolyn, Xanthines
Bipolar disorder	Lithium
Cardiac disease	Beta adrenergic blockers, Dopamine, Calcium channel blockers, Class I a antiarrhythmic, Class I c antiarrhythmics, Class III antiarrhythmic, Procainamide, Disopyramide, Quinidine, Vasodilator nitrates, Diuretic loops
Coronary/peripheral vascular disease	Antiplatelet, Oral anticoagulants, Trental
Cystic fibrosis	Anti-inflammatory Glucocorticoids, Enzymes
Depression	Monoamine oxidase inhibitors, Phenothiazine combinations, Tricyclic anti-depressants, SSRIs
Diabetes	Biguanides, Insulins, Sulfonylureas
Epilepsy	Anticonvulsants
End Stage Renal Disease	Marrow stimulants, Human erythropoietin
Gastric acid disorder	Histamine H2 blockers, Prostaglandins, Proton pump inhibitor
Gout	Colchicine, Uric acid inhibitors
HIV	Miscellaneous antiprotozoal, antivirals, pentamidine
Hyperlipidemia	Antilipemic clofibrate, Antilipidemic exchange resins, HMG coagulant reductase Inhibitors
Hypertension	Angiotensin Converting Enzyme (ACE) inhibitors, Beta adrenergic blockers, Dopamine, Calcium channel blockers, Antihypertensive vasodilators, Clonidine, Ganglionic blockers, Guanethidine, Methyldopa, Rauwolfia alkaloids, Alpha/Beta blockers, Diuretic combinations, Diuretic k depleting agents, Diuretic k sparing agents
Irritable bowel syndrome	Sulfonamides
Liver disease	Ammonia detoxicants

Malignancies	Leucovorin, Monoclonal, Miscellaneous antinauseants, Antineoplastic alkylating, Antineoplastic antibiotics, Antineoplastic MAO inhibitors, Antineoplastic progesterones, Antineoplastic pyrimidines, Antineoplastics misc, Bladder protectant, Methotrexate, Purine antimetabolites, Colony stimulating factors
Pain	Nonsteroidal anti-inflammatory agents
Pain and Inflammation	Opiates
Parkinsons disease	Dopamine, MAO b inhibitors
Psychotic illness	Miscellaneous antipsychotics, Butyrophenones, Phenothiazines, Thiothixenes
Renal disease	Potassium removing resins
Rheumatoid arthritis	Antiinflammatory Glucocorticoids, Gold salts-injectable, Gold salts-oral
Thyroid disorder	Thyroid replacement
Transplant	Immunosuppressive agents
Tuberculosis	Anti-tuberculosis antibiotics, Isoniazide

APPENDIX-G: ADDITIONAL RESULTS - ACEI/ARB MEASURE

TABLE A-5: Agreement in outliers: Crude and case-mix adjusted ACEI/ARB adherence scores

Groups Based on Crude Estimates	Groups Based on Case-mix Adjusted Estimates					
	Classical Logistic Regression Model			Hierarchical Logistic Regression Random Intercept Model		
	Low	Medium	High	Low	Medium	High
Low	11	6	6	14	7	2
Medium	2	0	1	2	1	0
High	0	0	17	0	2	15
Percentage misclassified[a]	15.4%	100.0%	29.2%	12.5%	90.0%	11.8%
Cohen's κ[b]	0.43			0.52		

[a]Percentage misclassified was calculated assuming risk adjustment method as the correct classification.
[b]Cohen's Kappa coefficient is a measure of agreement between unadjusted performance ratings and ratings based on the risk adjustment method with higher values indicating greater agreement

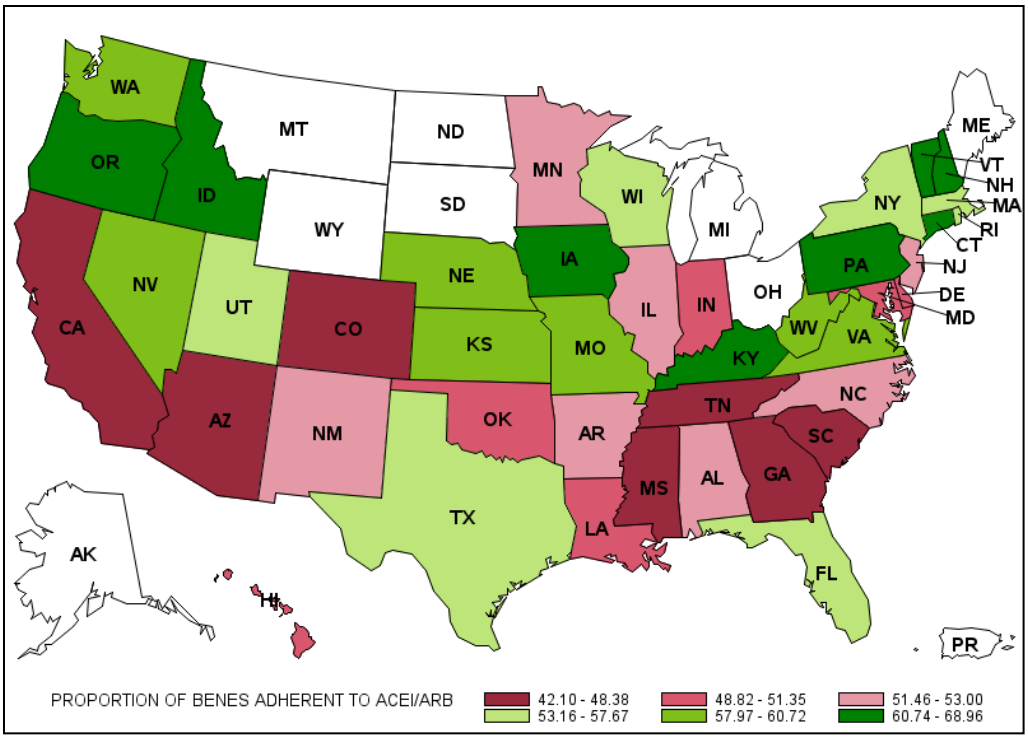

FIGURE A-3: Interstate variations in ACEI/ARB crude estimates

APPENDIX-H: ADDITIONAL RESULTS - BB MEASURE

TABLE A-6: Agreement in outliers: Crude and case-mix adjusted BB adherence scores

Groups Based on Crude Estimates	Groups Based on Case-mix Adjusted Estimates					
	Classical Logistic Regression Model			Hierarchical Logistic Regression Random Intercept Model		
	Low	Medium	High	Low	Medium	High
Low	11	4	2	13	3	1
Medium	1	5	2	2	5	1
High	0	1	17	0	1	17
Percentage misclassified[a]	8.3%	50.0%	19.0%	13.3%	44.4%	10.5%
Cohen's κ[b]	0.64			0.71		

[a]Percentage misclassified was calculated assuming risk adjustment method as the correct classification.
[b]Cohen's Kappa coefficient is a measure of agreement between unadjusted performance ratings and ratings based on the risk adjustment method with higher values indicating greater agreement

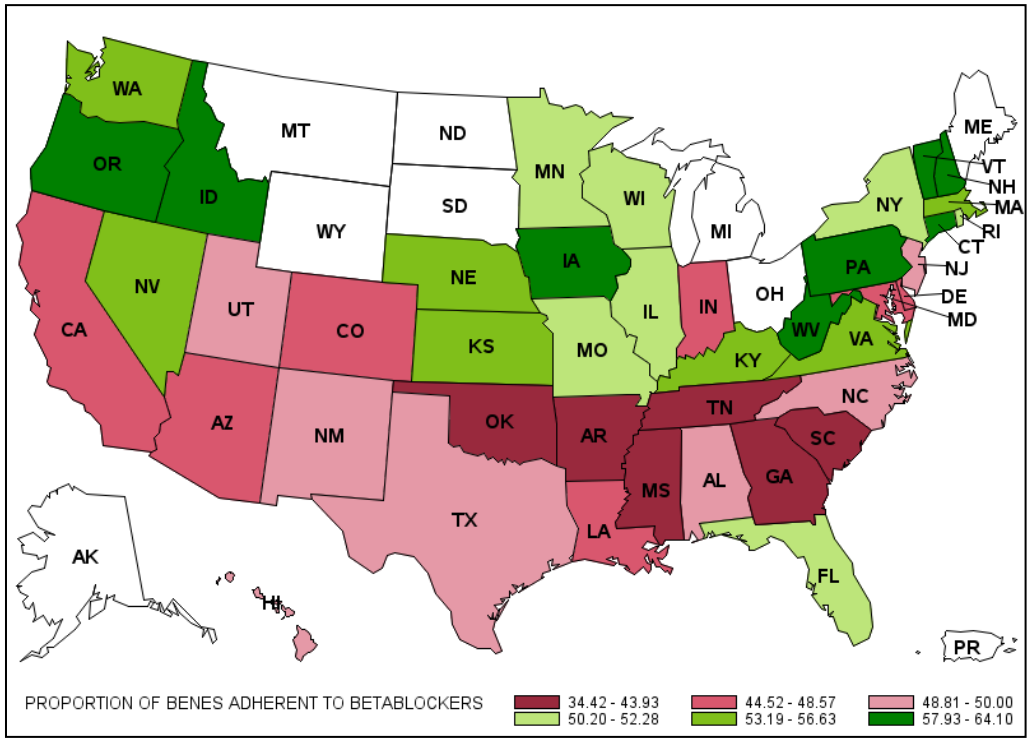

FIGURE A-4: Interstate variations in BB crude estimates

APPENDIX-I: ADDITIONAL RESULTS - CCB MEASURE

TABLE A-7: Agreement in outliers: Crude and case-mix adjusted CCB adherence scores

Groups Based on Crude Estimates	Groups Based on Case-mix Adjusted Estimates					
	Classical Logistic Regression Model			Hierarchical Logistic Regression Random Intercept Model		
	Low	Medium	High	Low	Medium	High
Low	10	7	2	15	3	1
Medium	0	5	2	1	4	2
High	0	2	15	0	4	13
Percentage misclassified[a]	0.0%	64.3%	21.1%	6.3%	63.6%	18.8%
Cohen's κ[b]	0.55			0.60		

[a]Percentage misclassified was calculated assuming risk adjustment method as the correct classification.
[b]Cohen's Kappa coefficient is a measure of agreement between unadjusted performance ratings and ratings based on the risk adjustment method with higher values indicating greater agreement

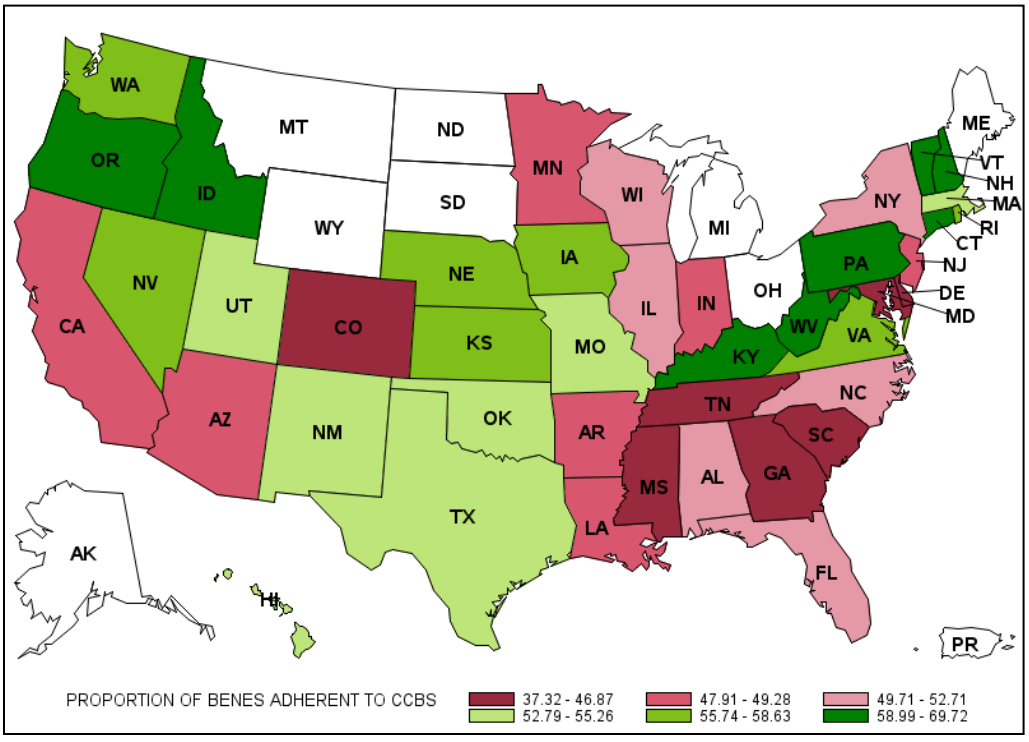

FIGURE A-5: Interstate variations in CCB crude estimates

APPENDIX-J: ADDITIONAL RESULTS - BIGU MEASURE

TABLE A-8: Agreement in outliers: Crude and case-mix adjusted BIGU adherence scores

Groups Based on Crude Estimates	Groups Based on Case-mix Adjusted Estimates					
	Classical Logistic Regression Model			Hierarchical Logistic Regression Random Intercept Model		
	Low	Medium	High	Low	Medium	High
Low	14	5	0	14	5	0
Medium	1	5	3	1	5	3
High	0	2	13	0	2	13
Percentage misclassified[a]	6.7%	58.3%	18.8%	6.7%	58.3%	18.8%
Cohen's κ[b]	0.61			0.61		

[a]Percentage misclassified was calculated assuming risk adjustment method as the correct classification.
[b]Cohen's Kappa coefficient is a measure of agreement between unadjusted performance ratings and ratings based on the risk adjustment method with higher values indicating greater agreement

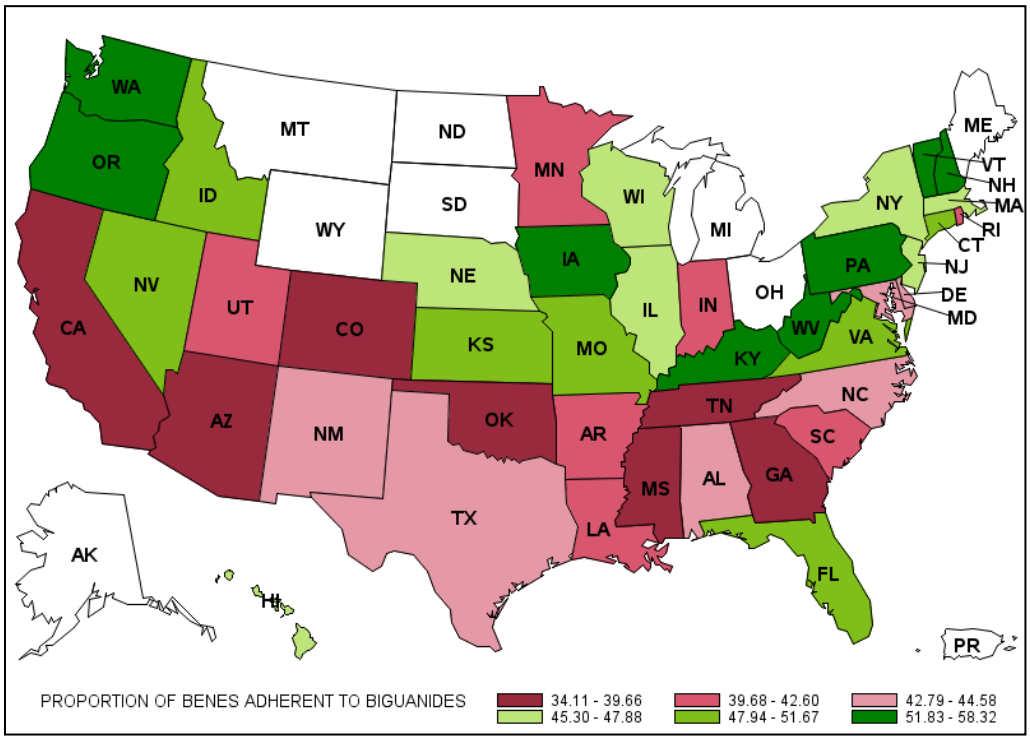

FIGURE A-6: Interstate variations in BIGU crude estimates

APPENDIX-K: ADDITIONAL RESULTS - SU MEASURE

TABLE A-9: Agreement in outliers: Crude and case-mix adjusted SU adherence scores

Groups Based on Crude Estimates	Groups Based on Case-mix Adjusted Estimates					
	Classical Logistic Regression Model			Hierarchical Logistic Regression Random Intercept Model		
	Low	Medium	High	Low	Medium	High
Low	11	4	0	13	2	0
Medium	2	8	4	2	10	2
High	0	4	10	0	5	9
Percentage misclassified[a]	15.4%	50.0%	28.6%	13.3%	41.2%	18.2%
Cohen's κ[b]	0.51			0.62		

[a]Percentage misclassified was calculated assuming risk adjustment method as the correct classification.
[b]Cohen's Kappa coefficient is a measure of agreement between unadjusted performance ratings and ratings based on the risk adjustment method with higher values indicating greater agreement

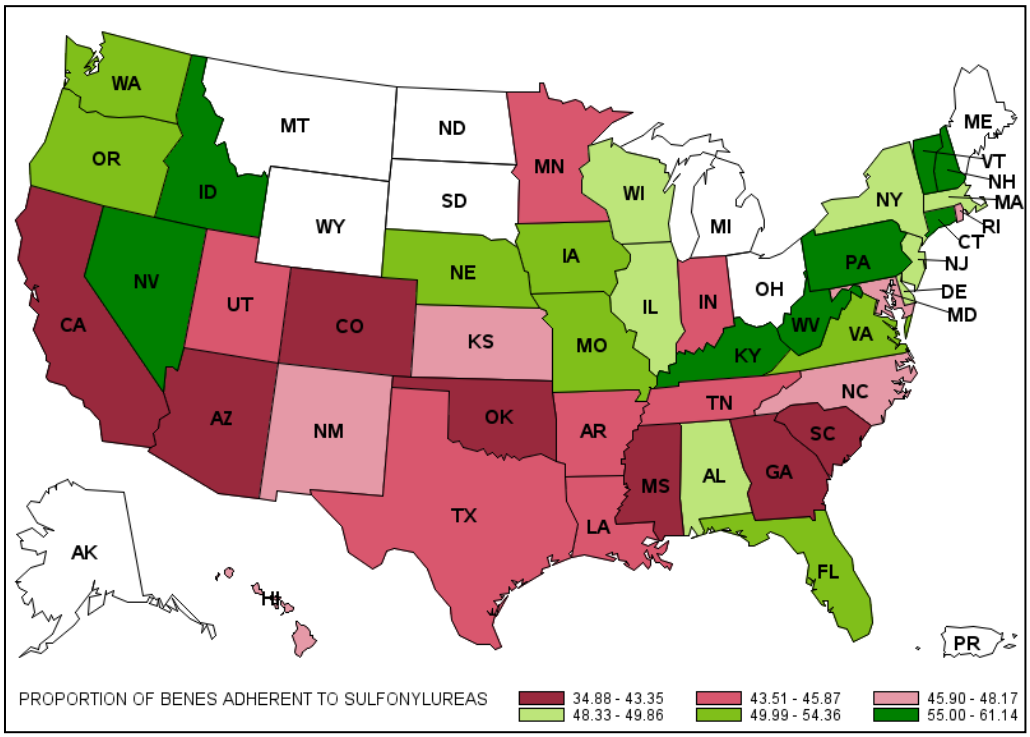

FIGURE A-7: Interstate variations in SU crude estimates

APPENDIX-L: ADDITIONAL RESULTS - TZD MEASURE

TABLE A-10: Agreement in outliers: Crude and case-mix adjusted TZD adherence scores

Groups Based on Crude Estimates	Groups Based on Case-mix Adjusted Estimates					
	Classical Logistic Regression Model			Hierarchical Logistic Regression Random Intercept Model		
	Low	Medium	High	Low	Medium	High
Low	11	4	2	13	3	1
Medium	1	9	2	1	10	1
High	0	2	12	0	2	12
Percentage misclassified[a]	8.3%	40.0%	25.0%	7.1%	33.3%	14.3%
Cohen's κ[b]	0.62			0.72		

[a]Percentage misclassified was calculated assuming risk adjustment method as the correct classification.
[b]Cohen's Kappa coefficient is a measure of agreement between unadjusted performance ratings and ratings based on the risk adjustment method with higher values indicating greater agreement

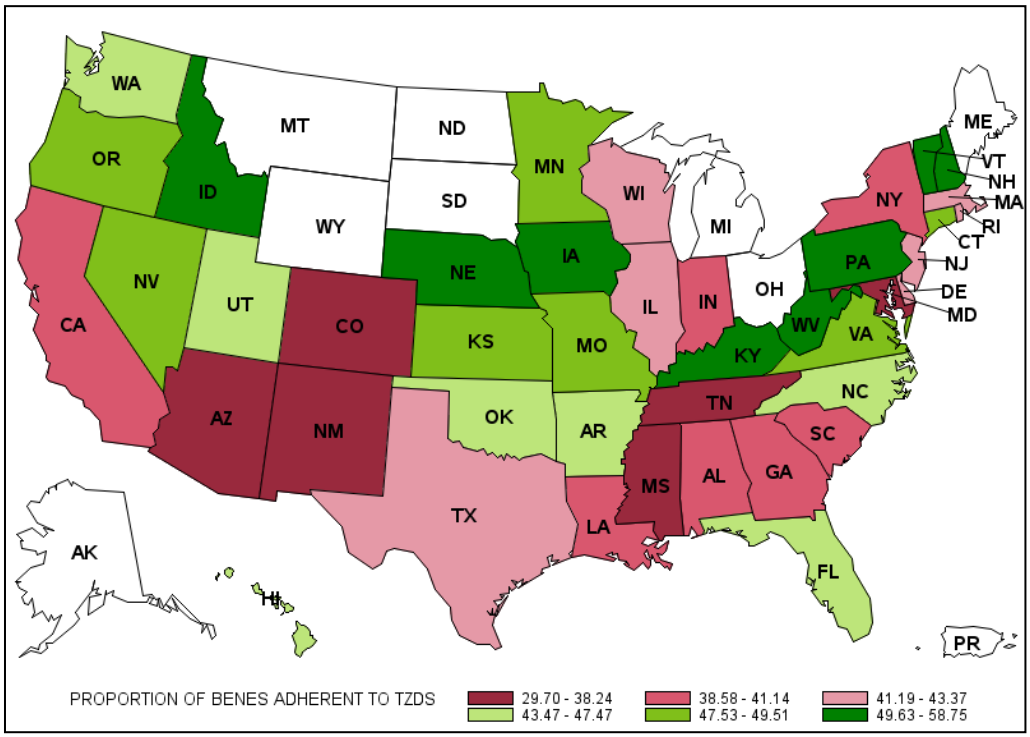

FIGURE A-8: Interstate variations in TZD crude estimates

APPENDIX-M: ADDITIONAL RESULTS - STAT MEASURE

TABLE A-11: Agreement in outliers: Crude and case-mix adjusted STAT adherence scores

Groups Based on Crude Estimates	Groups Based on Case-mix Adjusted Estimates					
	Classical Logistic Regression Model			Hierarchical Logistic Regression Random Intercept Model		
	Low	Medium	High	Low	Medium	High
Low	13	5	2	16	3	1
Medium	0	3	4	0	6	1
High	0	0	16	0	3	13
Percentage misclassified[a]	0.0%	62.5%	27.3%	0.0%	50.0%	13.3%
Cohen's κ[b]	0.60			0.71		

[a]Percentage misclassified was calculated assuming risk adjustment method as the correct classification.
[b]Cohen's Kappa coefficient is a measure of agreement between unadjusted performance ratings and ratings based on the risk adjustment method with higher values indicating greater agreement

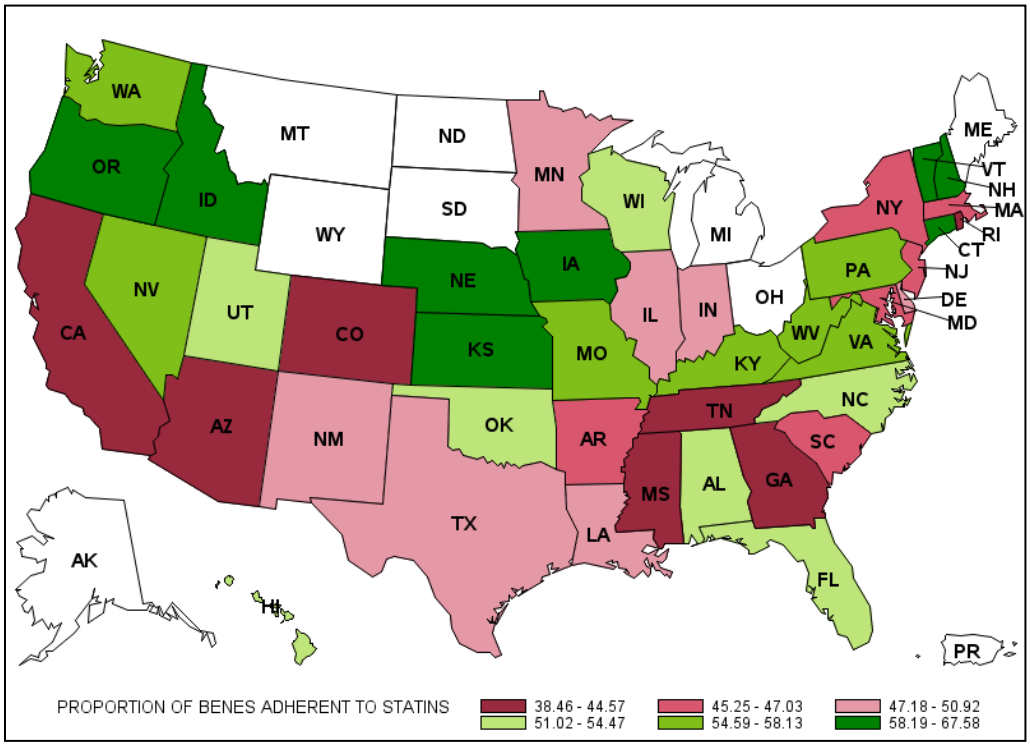

FIGURE A-9: Interstate variations in STAT crude estimates

APPENDIX-N: ADDITIONAL RESULTS - AD MEASURE

TABLE A-12: Agreement in outliers: Crude and case-mix adjusted AD measure scores

Groups Based on Crude Estimates	Groups Based on Case-mix Adjusted Estimates					
	Classical Logistic Regression Model			Hierarchical Logistic Regression Random Intercept Model		
AD Acute	Low	Medium	High	Low	Medium	High
Low	8	4	0	8	4	0
Medium	4	13	0	4	13	0
High	0	4	10	0	5	9
Percentage misclassified[a]	33.3%	38.1%	0.0%	33.3%	40.9%	0.0%
Cohen's κ[b]		0.57			0.54	
AD Chronic	Low	Medium	High	Low	Medium	High
Low	8	4	0	8	4	0
Medium	2	14	2	2	13	3
High	1	3	9	1	3	9
Percentage misclassified[a]	27.3%	33.3%	18.2%	27.3%	35.0%	25.0%
Cohen's κ[b]		0.57			0.53	

[a]Percentage misclassified was calculated assuming risk adjustment method as the correct classification.
[b]Cohen's Kappa coefficient is a measure of agreement between unadjusted performance ratings and ratings based on the risk adjustment method with higher values indicating greater agreement

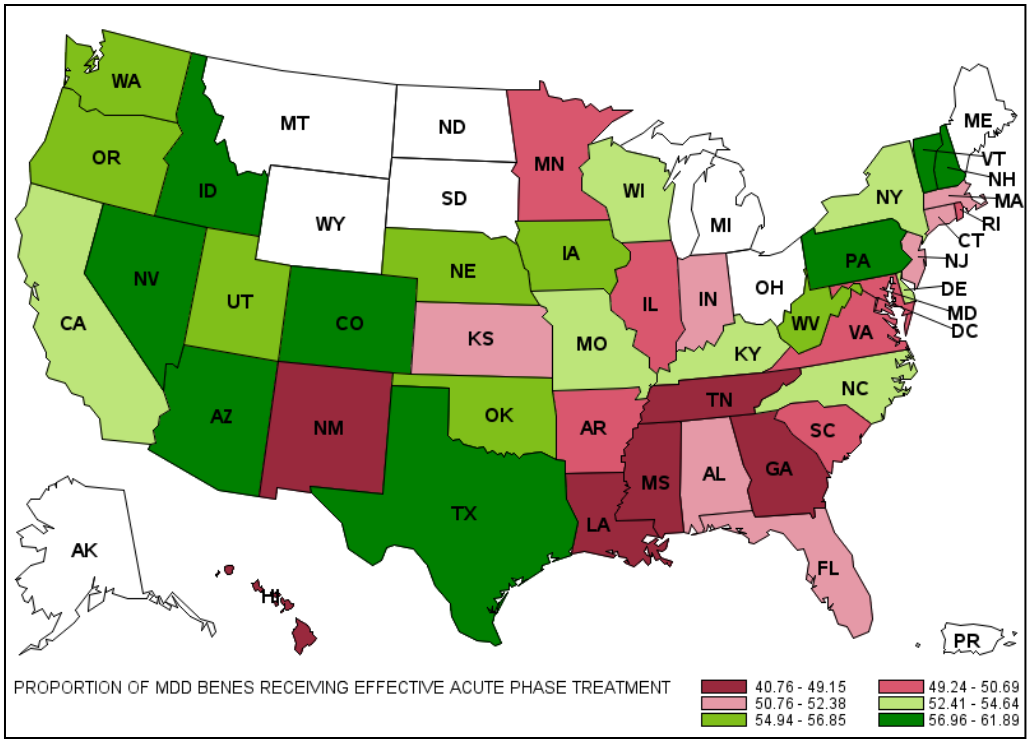

FIGURE A-10a: Interstate variations in AD Acute crude estimates

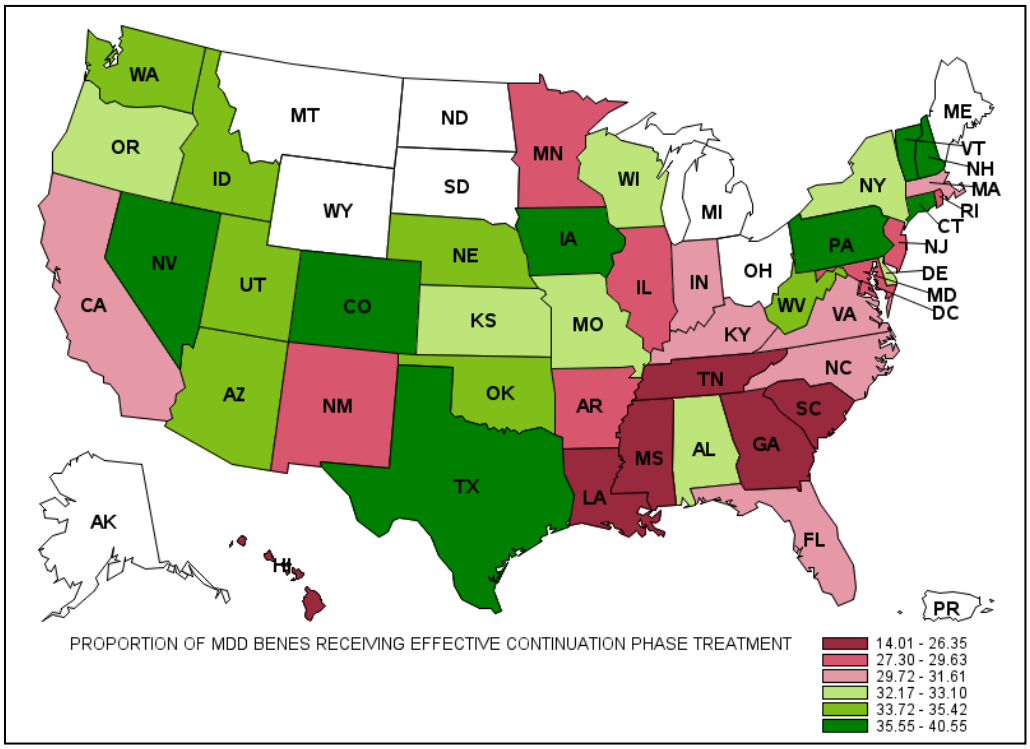

FIGURE A-11b: Interstate variations in AD Chronic crude estimates

APPENDIX-O: ADDITIONAL RESULTS - CAD MEASURE

TABLE A-13: Agreement in outliers: Crude and case-mix adjusted CAD measure scores

Groups Based on Crude Estimates	Groups Based on Case-mix Adjusted Estimates					
	Classical Logistic Regression Model			Hierarchical Logistic Regression Random Intercept Model		
	Low	Medium	High	Low	Medium	High
Low	9	3	2	9	3	2
Medium	5	5	2	5	5	2
High	0	7	10	0	7	10
Percentage misclassified[a]	35.7%	66.7%	28.6%	35.7%	66.7%	28.6%
Cohen's κ[b]	0.34			0.34		

[a]Percentage misclassified was calculated assuming risk adjustment method as the correct classification.
[b]Cohen's Kappa coefficient is a measure of agreement between unadjusted performance ratings and ratings based on the risk adjustment method with higher values indicating greater agreement

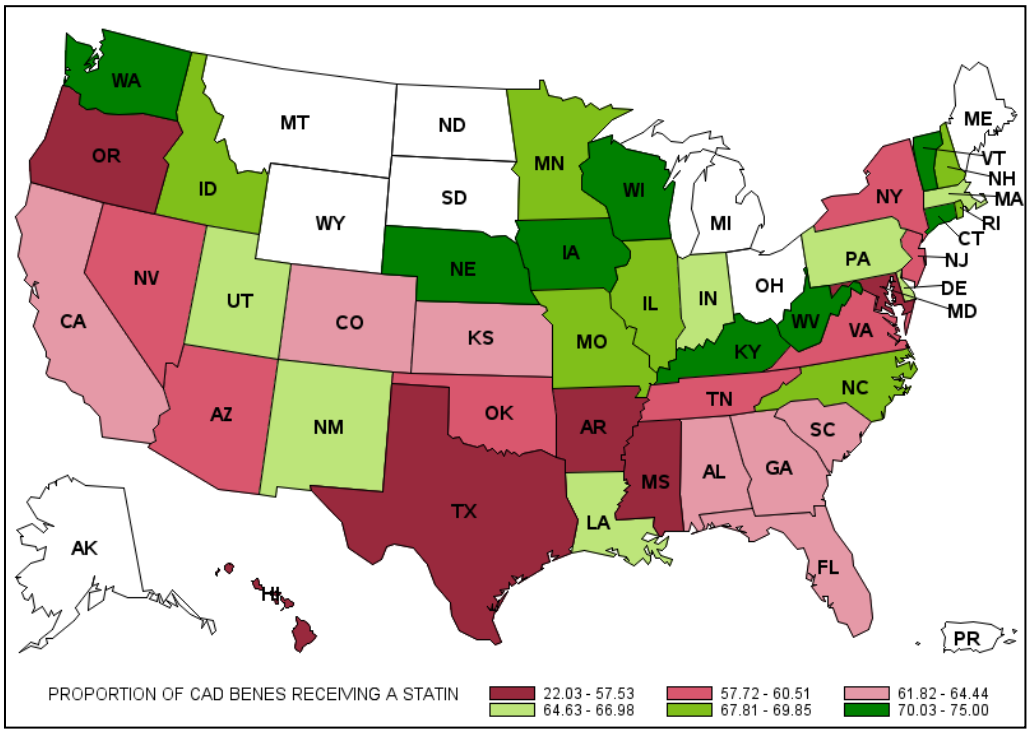

FIGURE A-12: Interstate variations in CAD crude estimates

APPENDIX-P: ADDITIONAL RESULTS - ICS MEASURE

TABLE A-14: Agreement in outliers: Crude and case-mix adjusted ICS measure scores

Groups Based on Crude Estimates	Groups Based on Case-mix Adjusted Estimates					
	Classical Logistic Regression Model			Hierarchical Logistic Regression Random Intercept Model		
	Low	Medium	High	Low	Medium	High
Low	14	0	0	14	0	0
Medium	2	11	0	1	12	0
High	0	2	14	0	0	16
Percentage misclassified[a]	12.5%	15.4%	0.0%	6.7%	0.0%	0.0%
Cohen's κ[b]	0.86			0.97		

[a]Percentage misclassified was calculated assuming risk adjustment method as the correct classification.
[b]Cohen's Kappa coefficient is a measure of agreement between unadjusted performance ratings and ratings based on the risk adjustment method with higher values indicating greater agreement

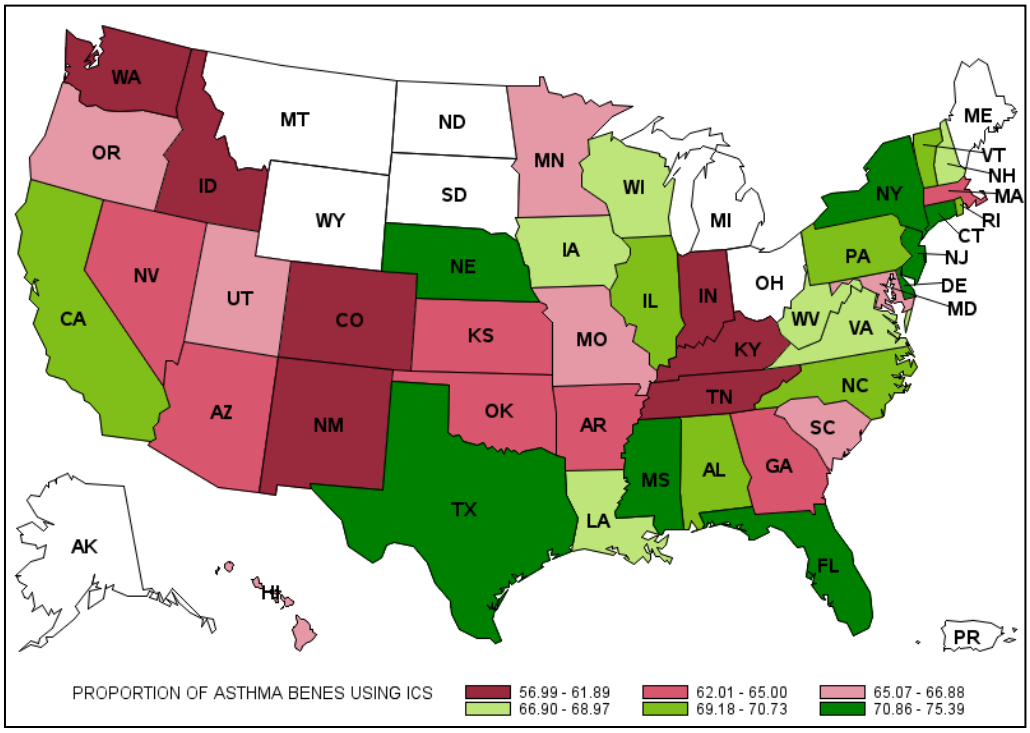

FIGURE A-13: Interstate variations in ICS crude estimates

APPENDIX-Q: ADDITIONAL RESULTS – MI1 MEASURE

TABLE A-15: Agreement in outliers: Crude and case-mix adjusted MI1 measure scores

Groups Based on Crude Estimates	Groups Based on Case-mix Adjusted Estimates					
	Classical Logistic Regression Model			Hierarchical Logistic Regression Random Intercept Model		
	Low	Medium	High	Low	Medium	High
Low	5	3	0	6	2	0
Medium	0	16	1	0	16	1
High	0	2	9	0	3	8
Percentage misclassified[a]	0.0%	23.8%	10.0%	0.0%	23.8%	11.1%
Cohen's κ[b]	0.73			0.73		

[a]Percentage misclassified was calculated assuming risk adjustment method as the correct classification.
[b]Cohen's Kappa coefficient is a measure of agreement between unadjusted performance ratings and ratings based on the risk adjustment method with higher values indicating greater agreement

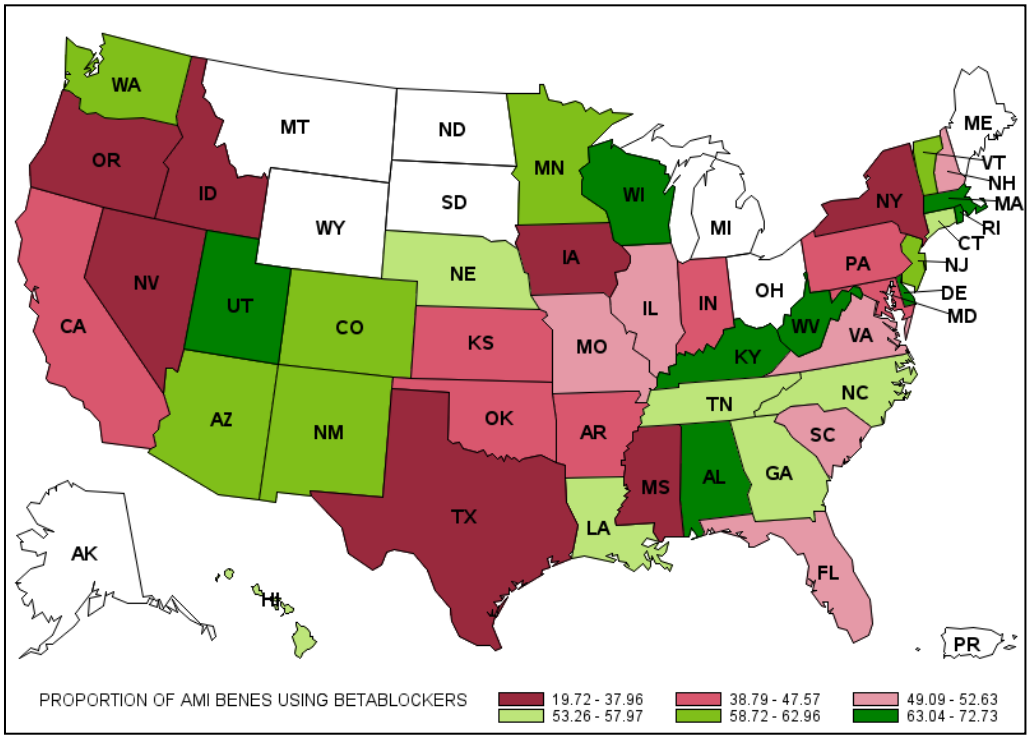

FIGURE A-14: Interstate variations in MI1 crude estimates

APPENDIX-R: ADDITIONAL RESULTS - MI MEASURE

TABLE A-16: Agreement in outliers: Crude and case-mix adjusted MI measure scores

Groups Based on Crude Estimates	Groups Based on Case-mix Adjusted Estimates					
	Classical Logistic Regression Model			Hierarchical Logistic Regression Random Intercept Model		
	Low	Medium	High	Low	Medium	High
Low	1	1	0	1	1	0
Medium	1	23	0	0	24	0
High	0	2	2	0	2	2
Percentage misclassified[a]	50.0%	11.5%	0.0%	0.0%	11.1%	0.0%
Cohen's κ[b]	0.55			0.63		

[a]Percentage misclassified was calculated assuming risk adjustment method as the correct classification.
[b]Cohen's Kappa coefficient is a measure of agreement between unadjusted performance ratings and ratings based on the risk adjustment method with higher values indicating greater agreement

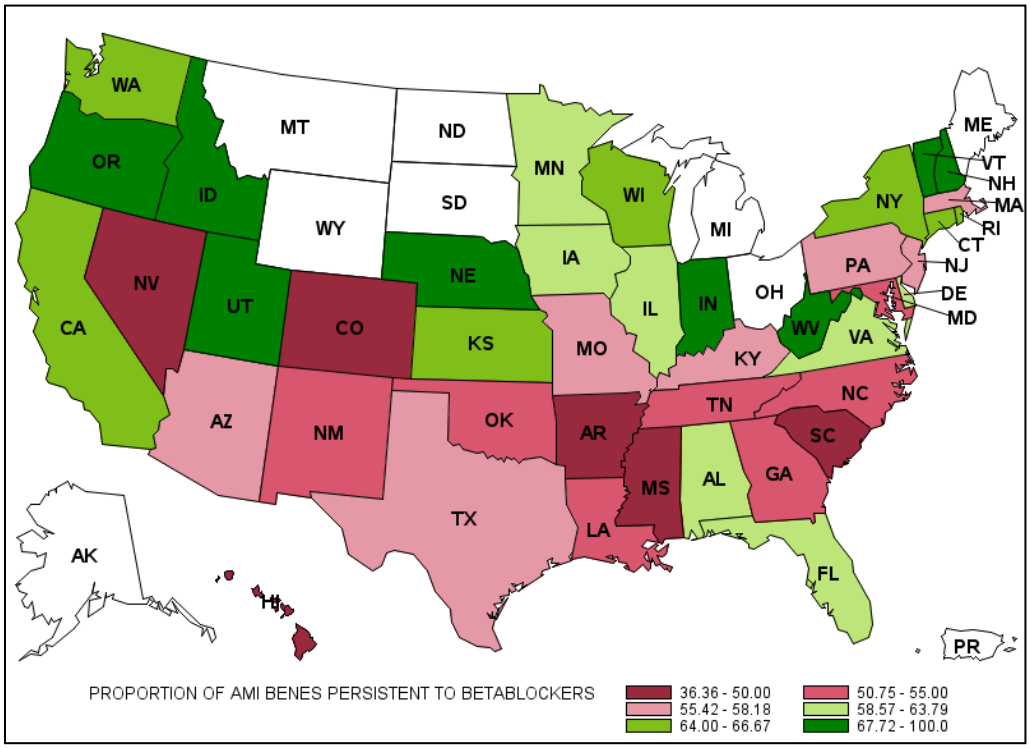

FIGURE A-14: Interstate variations in MI crude estimates

APPENDIX-S: ADDITIONAL RESULTS - COMPOSITE MEASURE

TABLE A-17: Agreement in outliers: Crude and case-mix adjusted composite measure scores

Groups Based on Crude Estimates	Groups Based on Case-mix Adjusted Estimates		
	Hierarchical Logistic Regression Three-Level Model		
	Low	Medium	High
Low	15	3	6
Medium	0	1	3
High	0	0	15
Percentage misclassified[a]	0.0%	75.0%	37.5%
Cohen's κ[b]	0.54		

[a]Percentage misclassified was calculated assuming risk adjustment method as the correct classification.

[b]Cohen's Kappa coefficient is a measure of agreement between unadjusted performance ratings and ratings based on the risk adjustment method with higher values indicating greater agreement.

VITA

Vennela Thumula was born in Hyderabad, India on August 12, 1984. She is the daughter of Dharma Rao Thumula and Saraswathi Ponugoti. She received her Bachelor's of Pharmacy (Hons.) from the Birla Institute of Technology & Sciences, Pilani in 2005. Vennela received her Master's in Pharmacy Administration from the University of Mississippi in 2010. After completion of her Master's degree, Vennela started working towards a Doctor of Philosophy degree in Pharmaceutical Sciences with an emphasis in pharmaceutical marketing and management. At the university she held research and teaching assistantship positions in the Department of Pharmacy Administration and received a research fellowship from the Center for Pharmaceutical Marketing and Management. She holds membership in Phi Kappa Phi Honor Society, Rho Chi Honor Society and other professional organizations like ISPOR, AcademyHealth and APHA. She served as the vice-president and president of ISPOR chapter at the University of Mississippi and as the chairperson of the survey committee of ISPOR student network. Vennela has a particular interest and expertise in health policy analysis, healthcare quality measurement, analyzing large claims data sets, examining disparities in healthcare/medication utilization, pharmaceutical product adoption, pricing and reimbursement issues. Vennela has done an internship with the Agency for Healthcare Research & Quality, Rockville, Maryland. After completion of her Ph.D. Vennela will be working as a policy analyst at the Worker's Compensation Research Institute, Cambridge, Massachussetts.